

Psychological Bulletin

HARRY HELSON, Editor
Kansas State University

CONTENTS

Biometrical Genetics and Behavior: Reanalysis of Published Data...	
.....P. L. BROADHURST AND J. L. JINKS	337
Teaching Machines: A Review.....	CHARLES S. MORRILL 363
Development of Research on the Physiological Mechanisms of Auditory Localization.....	MARK R. ROSENZWEIG 376
Physiological Effects of "Hypnosis"...	THEODORE KINOPHON BARBER 390
The Number Concept: A Phylogenetic Review.....	FRANK WESLEY 420
Hypnotic Age Regression.....	AUBREY J. YATES 429

Published Bimonthly by the
American Psychological Association

VOL. 53, No. 5

SEPTEMBER 1961

Consulting Editors

W. DRYAN, JR.
Kansas State University

R. R. BLAKE
University of Texas

W. R. GARNER
Johns Hopkins University

J. P. GUILFORD
University of Southern California

W. H. HOLSTENAU
University of Texas

L. J. POSTMAN
University of California, Berkeley

J. A. ROTTER
Ohio State University

S. B. SELLS
Texas Christian University

W. A. WILSON, JR.
Brigham Young College

The *Psychological Bulletin* contains evaluative reviews of research literature and reviews of research methodology and instrumentation in psychology. This JOURNAL does not publish reports of original research or original theoretical articles.

Manuscripts should be sent to the Editor, Harry Helson, Department of Psychology, Kansas State University, Manhattan, Kansas.

Preparation of articles for publication. Authors are strongly advised to follow the general directions given in the *Publication Manual of the American Psychological Association* (1957 Revision). Special attention should be given to the section on the preparation of the references (pp. 50-60), since this is a particular source of difficulty in long reviews of research literature. All copy must be double spaced, including the references. All manuscripts should be submitted in duplicate, one of which should be an original typed copy; author's name should appear only on title page. Dittoed and mimeographed copies are not acceptable and will not be considered. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail and to check carefully the typing of the final copy.

Reprints. Fifty free reprints are given to contributors of articles and notes.

ARTHUR C. HOFFMAN
Production Manager

HULEN ORR
Promotion Manager

Communications—including subscriptions, orders of back issues, and changes of address—should be addressed to the American Psychological Association, 1333 Sixteenth Street N.W., Washington 6, D.C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Annual subscription: \$10.00 (Foreign \$10.50). Single copies, \$2.00.

PUBLISHED BIMONTHLY BY:
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Menasha, Wisconsin
and 1333 Sixteenth Street N.W., Washington 6, D.C.

Second-class postage paid at Washington, D.C., and at additional mailing offices. Printed in U.S.A.
Copyright The American Psychological Association, Inc., 1961.

Psychological Bulletin

BIOMETRICAL GENETICS AND BEHAVIOR: REANALYSIS OF PUBLISHED DATA¹

P. L. BROADHURST

Institute of Psychiatry, University of London

AND J. L. JINKS²

ARC Unit of Biometrical Genetics, University of Birmingham

It is the purpose of this paper to scrutinize the attempts which have been made to provide quantitative data relating to the inheritance of behavioral characteristics in infra-human animals, and to reanalyze these data in terms of the polygenic or multifactorial hypothesis of genetical determination. Much of the data derived from the psychological field shows continuous variation and is consequently of the sort which lends itself to such polygenic analysis, as opposed to that employed in the analysis of discrete characteristics typical of classical Mendelian genetics. While it should be noted that there are now several experimental methods and analyses which have been developed for dealing with polygenic inheritance, it is not our present intention to undertake an evaluative survey of their relative merits as applied to psychogenetics at its current stage of development. Instead we propose to concentrate on and to employ one set of such tech-

niques, those of biometrical genetics as developed by Mather (1949) expressly for the analysis of continuous variation, especially in plants, and which we judge to be particularly promising in their application to the inheritance of behavior. An introduction to the general model and assumptions of this biometrical approach as applied to psychogenetics will be found in Broadhurst (1960).

EXPERIMENTAL METHOD

Few experiments in psychogenetics have been of a kind which can lead to a partitioning of the variation into its heritable and nonheritable components. Even fewer have been designed in such a way that the various tests are sensitive and the analysis reliable. Satisfactory experimental procedures for applying biometrical analysis in psychogenetics have recently been discussed by Broadhurst (1960) and we will merely note the following points: (a) the experimental animals should be randomized and the experiments replicated; (b) the parental stocks must be inbred; and (c) for investigating a cross between two inbred strains at least the two parental (P), first and second filial (F_1 and F_2) and backcross (B) generations should be reared.

¹ A summarized version of this paper was read to the Society for Experimental Biology in London in January 1961.

² Harkness Fellow of the Commonwealth Fund in the Division of Biology, California Institute of Technology, Pasadena, California, 1959-60.

The published data which come nearest to satisfying these requirements are those of Dawson (1932), Brody (1942), Goy and Jakway (1959), Jakway (1959), and Thompson and Fuller.⁹ Although no special precautions were taken by Dawson and by Brody to ensure the homozygosity of the parental strains the genetical differences between them were much greater than those within them; and insofar as our interest is in the differences between the parental strains, differences within them may be regarded as a further source of error variations, like the nonheritable differences with which they are confounded. The procedures appropriate to the analysis of these five sets of data are outlined in the next sections.

Analysis of Means

Following Hayman and Mather's (1955) and Jinks and Jones' (1958) extension of Mather (1949) we can write the generation means of a cross between two inbred strains in terms of six parameters: m , $[d]$, $[h]$, $[i]$, $[j]$, and $[l]$ which are the mean, additive, dominance, and three nonallelic, first-order interaction components between pairs of genes, respectively.

Thus:

$$\bar{P}_1 = m + [d] + [i] - \frac{1}{2}[j] + \frac{1}{4}[l]$$

$$\bar{P}_2 = m - [d] + [i] + \frac{1}{2}[j] + \frac{1}{4}[l]$$

$$\bar{F}_1 = m + [h] + \frac{1}{4}[l]$$

$$\bar{F}_2 = m + \frac{1}{2}[h]$$

$$\bar{B}_1 = m + \frac{1}{2}[d] + \frac{1}{2}[h] + \frac{1}{4}[i]$$

$$\bar{B}_2 = m - \frac{1}{2}[d] + \frac{1}{2}[h] + \frac{1}{4}[i]$$

Hence from these generation means we can estimate the heritable components as:

⁹ W. R. Thompson and J. L. Fuller, personal communication, 1959.

$$[d] = \bar{B}_1 - \bar{B}_2$$

$$[h] = \bar{F}_1 - 4\bar{F}_2 - \frac{1}{2}\bar{P}_1 - \frac{1}{2}\bar{P}_2 + 2\bar{B}_1 + 2\bar{B}_2$$

$$[i] = 2\bar{B}_1 + 2\bar{B}_2 - 4\bar{F}_2$$

$$[j] = 2\bar{B}_1 - \bar{P}_1 - 2\bar{B}_2 + \bar{P}_2$$

$$[l] = \bar{P}_1 + \bar{P}_2 + 2\bar{F}_1 + 4\bar{F}_2 - 4\bar{B}_1 - 4\bar{B}_2$$

and their sampling errors (SE²) as:

$$V_{[d]} = V_{\bar{B}_1} + V_{\bar{B}_2}$$

$$V_{[h]} = V_{\bar{F}_1} + 16V_{\bar{F}_2} + \frac{1}{4}V_{\bar{P}_1} + \frac{1}{4}V_{\bar{P}_2} + 4V_{\bar{B}_1} + 4V_{\bar{B}_2}$$

$$V_{[i]} = 4V_{\bar{B}_1} + 4V_{\bar{B}_2} + 16V_{\bar{F}_2}$$

$$V_{[j]} = 4V_{\bar{B}_1} + V_{\bar{P}_1} + 4V_{\bar{B}_2} + V_{\bar{P}_2}$$

$$V_{[l]} = V_{\bar{P}_1} + V_{\bar{P}_2} + 4V_{\bar{F}_1} + 16V_{\bar{F}_2} + 16V_{\bar{B}_1} + 16V_{\bar{B}_2}$$

The standard errors of the components can thus be obtained and tests of their significance by the customary methods applied.

If the gene effects are additive, that is, the genes are independent in their action, then the three components which estimate the effects of nonallelic interactions, $[i]$, $[j]$, and $[l]$, will be nonsignificant and the following identities known as *scaling tests* will hold within the limits of sampling error (Mather, 1949):

$$\text{Test A: } \bar{P}_1 + \bar{F}_1 - 2\bar{B}_1 = 0$$

$$\text{Test B: } \bar{P}_2 + \bar{F}_1 - 2\bar{B}_2 = 0$$

$$\text{Test C: } \bar{P}_1 + \bar{P}_2 + 2\bar{F}_1 - 4\bar{F}_2 = 0$$

A joint test of these three identities has been devised by Cavalli (1952). In this we estimate weighted least squares values for m , $[d]$, and $[h]$ from the generation means, assuming the absence of nonallelic interactions. The weights used are the reciprocals of the squared standard deviations of the generation means. The squared deviations of the ex-

pected and observed generation means are then a χ^2 with $(n-3)$ degrees of freedom, where n is the number of observed generation means.

If this χ^2 is nonsignificant then nonallelic interactions are absent and we can interpret directly the estimates of $[d]$ and $[h]$ obtained in the scaling test. The ratio of

$$\frac{[h]}{[d]} = \frac{h}{r_d \Sigma d}$$

is the so-called "potence ratio" (Wigan, 1944) and this is a measure of dominance only if the genes are associated in the parent lines, that is r_d (the degree of association) = ± 1 (Jinks & Jones, 1958) and all $[h]$ increments have the same sign. The potence ratio can theoretically take any value between zero and infinity. While a significant potence ratio indicates dominance of the individual genes predominantly in the same direction, zero potence does not necessarily indicate absence of dominance.

If the χ^2 from the joint scaling test is significant then nonallelic interactions are present and these can be analyzed by estimating $[i]$, $[j]$, and $[l]$ and testing their significance. A comparison of the signs of $[l]$ and $[h]$ will then tell us the type of nonallelic interaction involved. If their signs are the same then cooperative or complementary interaction between the genes predominates while if their signs differ competitive or duplicate interaction predominates (Jinks & Jones, 1958). The component $[j] = r_j \Sigma j$ on the other hand provides us with an indication of the distribution of the interacting genes in the parental lines. Thus with complete association $r_j = \pm 1$ and $[j]$ may have a significant value, but with maximum dispersion $r_j = 0$ and $[j]$ must be zero.

Scales

Although allowance can be made for nonallelic interactions and genotype-environmental interactions in the analysis of second degree statistics (Hayman & Mather, 1955; Mather & Jones, 1958), with the paucity of second degree statistics available in the psychogenetical experiments to be analyzed we can merely attempt to find an empirical scale on which these effects make no significant contribution to the variation.

Clearly, our criterion of an adequate scale which eliminates nonallelic interaction is one which leads to a nonsignificant χ^2 in the joint scaling test. However, a scale which is empirically adequate for this purpose will not necessarily remove any genotype-environmental interaction which may be present, that is, lead to homogeneity of the variances of the parents and F_{1s} . We must, therefore, adopt a scale which at least minimizes and balances these two sources of bias.

Analysis of Variances into Components of Variation

On an adequate scale the variances (s^2) of the parent, F_1 , F_2 , and back-cross generations are (Mather, 1949):

$$V_{P_1} = V_{P_2} = V_{F_1} = E_1$$

$$V_{F_2} = \frac{1}{2}D + \frac{1}{4}H + E_1$$

$$V_{B_1} + V_{B_2} = \frac{1}{2}D + \frac{1}{2}H + 2E_1$$

where $D = \Sigma d^2$, $H = \Sigma h^2$ and E_1 are the additive, dominance, and non-heritable components of variation, respectively.

In addition we have

$$V_{B_1} - V_{B_2} = \pm \Sigma (dh)$$

Solution of these equations leads to estimates of D , H , and E_1 , from which

we can obtain estimates of dominance and heritability. The dominance ratio, H/D , will be zero for no dominance, one for complete dominance and greater than one for overdominance (heterosis). Heritability can be assessed in a variety of ways of which we will use two, $D/(D+E_1)$, i.e., the ratio of the additive variation to the sum of the additive plus nonheritable variation, and $(\frac{1}{2}D + \frac{1}{4}H)/(\frac{1}{2}D + \frac{1}{4}H + E_1)$, which is the proportion of heritable variation in an F_2 population. These ratios therefore represent estimates of heritability "in the narrow sense" and "in the broad sense," respectively.

When $\Sigma(dh)$ does not equal zero it supplies additional evidence for the presence of dominance. It also shows which parent carries the preponderance of dominant allelomorphs, for the backcross to this parent has the lower variance.

Number of Effective Factors

Only one estimate of the number of effective factors is applicable to the type of data so far obtained in psychogenetics, namely, the estimate of K , (Mather, 1949). This equals $\frac{1}{4}(\bar{P}_1 - \bar{P}_2)^2/D$ which for k genes of equal effect and associated in the parental lines equals k^2d^2/kd^2 .

In practice this estimate is always minimal because it assumes that the genes are associated (i.e., $r_d = 1$) and that all genes give equal increments (i.e., $d_a = d_b = d_c \dots$). It is, however, worth obtaining in the psychogenetical experiments because of the practice of deliberately selecting the most extreme lines available as parents in the cross. Such selection will lead to a preponderance of association in the parental lines, thus partially satisfying one of the assumptions required.

EXAMPLES FROM THE LITERATURE

We can now illustrate the analyses described earlier by reference to particular experiments in the field of psychogenetics.

Dawson

Dawson's work (1932) was most accomplished genetically, and still might serve as a model of how a psychogenetical investigation could be approached, at least, from the genetical aspect. Unfortunately, the more purely psychological treatment is not of comparable quality and leaves much to be desired. Dawson investigated the inheritance of wildness in mice, defining wildness in terms of the speed the animals showed in running down a straight runway. We shall give his description of the method.

The method of testing consisted in placing the mouse at one end of a runway and allowing it to run to the other end. The time required was recorded by means of a stopwatch. The runway was 24 feet long, 9½ inches wide and 13 inches high. The sides and ends were of galvanized sheet iron, the floor of soft wood. One foot from each end a black line was painted on the floor of the runway. The time required for the mouse to run from one line to the other, a distance of 22 feet, was recorded. A movable partition made of wall-board and bound with rubber was used to prevent the mouse from running back during the test and to aid in starting the test and capturing the mouse afterwards. The following procedure was carried out in testing the mice in this device. On the date that the mice in a certain pen were to be tested they were carried to the runway a short distance away and tested one at a time. The mouse to be tested was confined by the movable partition in a space about one foot from the end of the runway until everything was ready when the partition was raised and the stopwatch started as soon as the mouse crossed the black line. The mouse was followed by the experimenter with the partition which was placed in position to prevent the animal from running back if it showed any signs of doing so. Nothing was done to frighten the mouse other than the procedure described. This was usually

sufficient to cause even the tame mice to run or walk towards the other end of the runway. If the mouse ran swiftly, it was impossible to keep up with it with the partition; but if more slowly the partition was moved along and kept about twelve to fifteen inches behind the mouse. If the mouse stopped and showed no inclination to go forward the partition was slowly advanced until it touched the mouse. In all but three or four cases this was sufficient to start the mouse again. The few individuals where this was not the case were shoved a little and thus started. When the mouse crossed the line at the far end of the runway, the watch was stopped and the partition taken out of the runway allowing the mouse to run back to the starting point or in case it did not do so voluntarily it was urged by means of the partition. This prevented the mice from associating the far end of the runway with being caught. Since each individual was tested three times this point was of considerable importance. After the mouse had been cornered at the starting point by means of the partition, it was caught and the number in its ear read. . . . These trials were conducted at weekly intervals after the mouse reached 75 days of age. In order to facilitate the testing and caring for the mice, a variation of one day in either direction was permitted. Thus the first trial for a given mouse might occur on the 74th, 75th or 76th day. A few trials had to be made on different dates. The trials were in nearly all cases made in the evening or at night when there was very little outside disturbance to distract the mice. The lighting was kept as far as possible the same throughout the experiment (pp. 299-300).

It will be seen that a large subjective element could enter into the factors determining the speed of running of a given subject in the alley through the way in which the partition was manipulated. This is not entirely overcome by the procedure of identifying the animal after completing the test, as, in the parental generations at any rate, there were distinctive coat color differences between the strains. Nevertheless, the corrected reliability coefficient for the three trials for all the 1,232 subjects used is reported as $0.92 \pm (SE)0.04$.

The subjects were a strain of wild

mice which had been reared in the laboratory for several years and three strains of tame mice, an albino, and two strains of brown mice with pink eyes, one of them also having short ears. The wild strain of mice appeared to be more highly inbred than the tame mice since selection among the former produced no response while the latter responded. This, however, as we have seen earlier, is not a serious problem in view of the large and high significant differences between the wild and tame strains for the measure under consideration.

Dawson was able to extract a large amount of information from his results regarding the nature of the genetical control of behavior in his runway situation. He showed that there was no linkage with sex or with any of the major gene effects identifiable in his strains, and, by reciprocal crossing, that there were probably no directional maternal effects (Broadhurst, 1961). He concluded that the wild-type behavior was dominant and that only a few genes are involved in determining the reaction. This last conclusion is principally based on the result of fitting curves, derived from Mendelian ratios and assuming various numbers of genes up to three, to the observed distributions. However, he admitted that probably a number of modifying genes were also involved. Implicit in his estimate of the number of genes were assumptions concerning size of individual gene effects, the distribution of genes in the parental lines and their dominance relations. A biometrical analysis along the lines proposed here would, therefore, appear to be appropriate.

In Table 1 will be found the relevant generation means and their standard errors calculated from the data given by Dawson. The three

TABLE 1

DAWSON'S DATA: MEANS AND THEIR STANDARD ERRORS IN SECONDS AND (n)

Generation Means							
	F ₁ (Wild)	F ₁ (Tame)	F ₁	F ₂	B ₁	B ₂	
Males	6.7 ± 0.3 (43)	24.5 ± 1.0 (53)	7.6 ± 0.3 (76)	13.0 ± 0.6 (175)	6.6 ± 0.3 (26)	27.4 ± 3.9 (54)	20.8 ± 1.6 (50)
Females	5.3 ± 0.3 (47)	25.3 ± 1.2 (54)	6.9 ± 0.3 (58)	11.8 ± 0.5 (190)	6.2 ± 0.5 (24)	18.7 ± 1.5 (48)	
Both	5.9 ± 0.2 (90)	24.9 ± 0.8 (117)	7.2 ± 0.2 (164)	12.4 ± 0.4 (365)	6.4 ± 0.4 (50)	23.3 ± 2.2 (102)	19.7 ± 1.4 (98)

strains of tame mice were used in these crosses and the results pooled.

Apart from the backcross to the slower parent (B₂), the sexes are in good agreement and this failure in the backcross can be traced to four males whose scores were greater than 90. Individuals with such high scores are met with nowhere else in Dawson's experiments which included some second backcrosses, F₂s and F₃s in addition to the data given in the table. The results omitting the four males are also shown in Table 1. Their omission improves the already good agreement between sexes and the analyses can now be carried out on the pooled sexes. We may add that omitting these four individuals does not affect the interpretation of the data since the sex difference they suggest is borne out neither by Dawson's nor our own detailed analyses.

The joint scaling test (Cavalli, 1952) gives the following weighted least squares estimates from the pooled sexes,

$$m = 15.99, [d] = 10.10 \text{ and } [h] = -8.74$$

which when compared with the observed generation means give a $\chi^2_{(3)}$ of 9.4 ($p = 0.05 - 0.02$). There is therefore some nonallelic interaction present. However, its magnitude would not normally warrant rescaling but these data also show significant geno-type-environmental interaction ($p < 0.01$) on the linear scale, as determined by Pearson and Hartley's

test (1958) for inhomogeneity of the P₁, P₂, and F₁ variances. Two transformations have, therefore, been tried, a square root and a log transformation. Of these the latter was the more satisfactory and the joint scaling test repeated on the new scale gave $m = 1.197 \pm 0.032$, $[d] = 0.340 \pm 0.030$, and $[h] = -0.222 \pm 0.059$ which gave a satisfactory fit with the observed data ($\chi^2_{(3)} = 0.34$).

The same scalar change also removed the genotype-environmental interaction; hence a solution for D, H, and E₁ was attempted from the second degree statistics. These gave values⁴ of

$$D = 0.052 \pm 0.024$$

$$H = -0.008 \pm 0.032$$

$$E_1 = 0.020 \pm 0.005$$

$$\Sigma(dh) = -0.032$$

This gives a heritability estimated as $D/(D+E)$ of 72% and estimated as $(\frac{1}{2}D + \frac{1}{4}H)/(\frac{1}{2}D + \frac{1}{4}H + E)$ of 55%. The confidence limits ($p = .05$) for the first estimate of heritability are 61% and 83%. A further estimate of heritability can be extracted from Dawson's data. He assortatively mated his F₂ individuals to raise an F₃ generation and from the results

⁴ This is the only place in this paper where the data permit the estimate of standard errors for these components. No further suitable replication of observations, e.g., by the provision of the raw data for both sexes, as in this case, is encountered.

we can estimate the parent/offspring correlation. This turns out to be 0.51.

Our estimate of the minimal number of effective factors is 2.2 with confidence limits of 3.5 and 1.6. This is in good agreement with Dawson's estimate which as we have seen is also minimal.

Thus, the behavioral difference between the wild and tame mice investigated by Dawson is controlled by at least three effective factors whose contributions are additive and independent of the environment on a logarithmic scale but which interact with one another and with the environment on a linear scale. Estimates of [i], [j], and [l] on the linear scale show that [j] which equals 7.6 ± 2.9 is responsible for the non-allelic interactions. The genes have a significant additive and dominance effect although the latter is not apparent in the second degree statistics presumably due to the effect of sampling variation on the negative correlation between the estimates of D, H, and E₁. The potency ratio is negative and greater than zero which means that there is a preponderance of dominant genes in the low scoring, i.e., wild type, parent. The significant estimate of $\Sigma(dh)$ confirms this and also shows the presence of a dominance component of variation. Heritability is quite high and estimates from different sources give consistent results.⁸

Brody

The analysis of Brody's experi-

⁸ Since this paper was submitted for publication, estimates of the number of genes and heritability in the F₂, derived from another reanalysis of Dawson's data, have been published by Fuller and Thompson (1960). They used methods of analysis (Wright, 1952) similar to those proposed here, and their estimates are in substantial agreement with our own.

ments (1942) follows much the same pattern. She investigated the inheritance of voluntary cage activity in rats using the high and low selections for activity begun by Rundquist (1933). The number of revolutions of the activity cages, during the last 15 of a 21-day period was taken as the measure. Each rat was housed in these cages at some time between 60 and 100 days of age. Some measure of inbreeding was practiced from the fifth generation of selection although neither its degree nor the precautions taken to control environmental variation are stated. Our present concern is with the crosses made using the inactive and active strains at the twenty-first generation of selection. A complete program of breeding F₁, F₂, and backcrosses, B₁ and B₂, was carried out, and, moreover, repeated using these two strains at the twenty-second generation of selection as parents. In each case the strains were crossed reciprocally to give the F₁s.

Brody's results show reasonably good agreement between the values obtained in the two replications of her crossing program. An analysis of variance of the complete data gave significance for only two items: the difference between sexes, and the difference between generations, i.e., P₁, P₂, F₁, F₂, B₁ and B₂. There were no significant differences between the replicate crossing programs initiated at the twenty-first and the twenty-second generations of selection and no significant interactions between the three main effects. We can, therefore, pool the two sets of crosses and sexes for the biometrical analyses.

The joint scaling test on the pooled data given in Table 2 gave weighted least squares estimates of $m = 74.71 \pm 5.55$, $[d] = 63.03 \pm 5.56$ and $[h] = -3.41 \pm 9.32$. These did not pro-

TABLE 2

BRODY'S DATA: MEANS AND THEIR STANDARD ERRORS IN REVOLUTIONS $\times 10^{-3}$ AND (*n*)

Pooled Generation Means					
P ₁ (Inactive)	P ₂ (Active)	F ₁	F ₂	B ₁	B ₂
13.2 \pm 4.1 (136)	136.4 \pm 12.5 (67)	73.6 \pm 7.1 (193)	71.5 \pm 9.3 (260)	26.8 \pm 9.3 (79)	115.2 \pm 10.2 (136)

vide a satisfactory fit with the observed generation means, $\chi^2_{(3)} = 15.27$ ($p = 0.01 - 0.001$). Clearly genic interaction is present on this scale. Since these data are not published in a form amenable to rescaling we cannot attempt to remove the interaction; we can, however, investigate its nature.

Estimates of the components of the generation means and their standard errors showed that only *m*, [*d*], and [*j*] were significant hence weighted least squares estimates of these three components were made assuming the other components were zero. These estimates of $m = 72.14 \pm 3.61$, [*d*] = 81.54 ± 13.62 and [*j*] = 44.34 ± 28.90 provided a satisfactory fit with the observed generation means, $\chi^2_{(3)} = 5.28$ ($p = 0.20 - 0.10$).

A further argument in favor of rescaling is provided by the second degree statistics which show significant ($p = 0.01 - 0.001$) genotype-environmental interaction. Since, however, rescaling is impossible we must be cautious in interpreting the components of variation because of the possible bias from the [*j*]-type non-allelic interaction and the genotype-environmental interactions. *H* proved to be nonsignificant and negative therefore *D* and *E* were recalculated assuming *H* = 0 with the following result: *D* = 3410.72 and *E* = 1945.91.⁶ These values give 64%

and 44% as our two estimates of heritability and 1.95 as the minimal number of effective factors. The uncertainty of the latter estimate makes it impossible decisively to rule out Brody's own interpretation based on a single gene difference between her selected parents. On the other hand we can discount a single gene interpretation on the basis of the significant genic interaction.

Thus the difference between the spontaneous cage activity of the selected strains measured on a linear scale depends on at least two interacting genes which also interact with the environment. As in Dawson's experiment it is the [*j*]-type interaction which is responsible for the genic interaction. There is no evidence of dominance and the potency ratio is zero. It is possible, however, that dominant and recessive alleles are equally frequent in the two selected strains. Heritability is about the same as in Dawson's experiment but since the scale on which it is measured is unsatisfactory we cannot place too much reliance on its absolute magnitude.

An unusual feature of Brody's results requires further comment. Both her *F*₁s were made reciprocally. Both of them show the same paternal effect, that is, in the direction of a negative influence of the mother—mothers from the active strain tending

⁶ Despite Brody's replication, no estimate of standard errors can be given for these components which is not potentially subject to serious inflation due to the various differences

between the variances noted. The impossibility of rescaling therefore renders the replicates unsatisfactory as a source of an estimate of error variation.

to have offspring lower in activity than those from inactive mothers. Pearson and Hartley's (1958) exact test for homogeneity of variance showed only the data from the twenty-first generation crosses were suitable for analysis of variance, which was applied following Snedecor's method (1956) for dealing with unequal numbers in subgroups in a two-way classification. This analysis revealed an interaction between sex and strain as shown in Table 3, which may be summarized by saying that the significant tendency of the F_1 to be unlike the mother in respect of activity was more pronounced in the case of sons than of daughters. Sex linkage or paternal inheritance may be responsible for this complex situation, but to distinguish between the two would require a much more involved experimental design than that used by Brody.

Thompson and Fuller

The systematic program of research in psychogenetics which has been proceeding at the Roscoe B. Jackson Memorial Laboratory at Bar Harbor, Maine, for the last decade, has, as might be expected, produced work of high quality. Only one set of data, however, has so far become available⁷ which lends itself to the complete biometrical analysis proposed in this paper. This is the work of Thompson and Fuller (Fuller & Thompson, 1960, pp. 267-269; Thompson 1953, 1956, see Footnote 3).

Thompson and Fuller employed the two inbred strains of mice showing extremes of high and low activity from a previous study in which a total of 15 strains had been studied, and

⁷ We are indebted to W. R. Thompson for making a draft copy of the MS containing these data available to us prior to publication.

TABLE 3
BRODY'S DATA: MEANS OF RECIPROCAL
CROSSES IN REVOLUTIONS $\times 10^{-3}$
AND (n)

F ₁ Offspring		
Strain of Mother	Males	Females
Active	33.9 (7)	109.4 (24)
Inactive	87.9 (21)	120.7 (30)

tested a large number of subjects on each of two tests which they describe as follows:

[Test 1] consisted of an open-field 30 by 30 inches with walls $3\frac{1}{2}$ inches high, and a hinged wire-mesh top. The floor was painted gray and the walls a flat black. The floor was divided by lines into a grid of 36 squares, each 5 by 5 inches. At the base of every other square was placed a barrier, 5 by $3\frac{1}{2}$ by $3\frac{1}{2}$ inches, painted a flat black. Leading into the open-field at one corner was a starting box with a separate hinged top. Test 2 was a Y-maze with arms $11\frac{1}{2}$ inches long by 3 inches wide by $3\frac{1}{2}$ inches deep. Angles between arms were equal. One arm was painted black inside, another gray and the third white. The maze was covered by a removable wire-mesh top. An animal was started at the end of the gray arm farthest from the junction point. Observation of animals in both tests were made under dim illumination as follows: in Test 1, a record was made by a mechanical counter of the number of lines crossed by a mouse in a 10 minute period. In Test 2, a count was made of the number of half-arm units traversed during each of six 100 second periods. . . . The correlation between the two tests was approximately 0.60 (Thompson & Fuller, see Footnote 3).

The two parental strains are known to be highly inbred, but the measures taken to control environmental variation are not specified. The F_1 , F_2 , and backcrosses were bred from the two strains and given both tests. The possibility of order effects of one test upon the other in the resulting data is not discussed. Reciprocal crosses were made, and the results pooled,

as were those of the two sexes as shown in Table 4. Thompson and Fuller subjected the data from the first test to a square root transformation in order to equalize the variance of the parental and F_1 generations which it did successfully. The transformed data is included in Table 4.

The joint scaling test for Test 1, Test 2, and the transformed Test 1 data gave the following weighted least squares estimates:

Component	Test 1	Test 1 (transformed)	Test 2
m	267.6 \pm 10.9	12.50 \pm 0.41	166.1 \pm 4.1
[d]	257.8 \pm 10.5	9.98 \pm 0.39	82.6 \pm 4.0
[h]	29.2 \pm 20.5	5.13 \pm 0.76	-2.0 \pm 7.7
χ^2	24.3 ($p < 0.001$)	6.6 ($p = 0.10-0.05$)	16.9 ($p < 0.001$)

On the linear scales both Test 1 and Test 2 show unsatisfactory fits with additive gene action. Hence we have nonallelic interactions present. Analysis of these nonallelic interactions shows that the [j]-type interaction is again largely responsible for the failure of the linear scale, having values of 25.6 ± 8.3 in Test 1 and 55.8 ± 9.0 in Test 2. Unfortunately the data given by Thompson and Fuller do not suffice to allow rescaling. Analysis of their square root transformed data for Test 1 has, however, shown that rescaling can reduce the genic interaction. Since this scale also removes the significant genotype-environmental interaction in Test 1 we can confidently analyze the second degree statistics on the square root scale.

Only Test 1 on the linear scale, however, provides estimates of the components of variation D, H, and E_1 which are sensible, that is to say, give positive values for D: the other two tests do not. This result is not unexpected for Test 2 where interactions are present but it is unexpected for the square root transformed data of Test 1 where the in-

teractions are largely scaled out. We will, therefore, merely give estimates of the heritabilities for the F_2 population since this does not require the partitioning of the heritable components of variation (see above). For Test 1 the values are 73% and 53% for the linear and square root scales, respectively, and for Test 2, 26%. The low value in the latter test provides an additional reason for the failure to obtain sensible estimates of

the additive and dominance components of variation. With no estimate of D we cannot evaluate our estimate of the number of effective factors.

Thus the genes controlling the two behavior patterns in mice investigated by Thompson and Fuller have a large additive effect but show no dominance on the linear scale. They do, however, interact with one another and with the environment in a way which can be largely removed by a square root transformation. Once again it is the [j]-type of interaction which is mainly responsible for the failure of the linear scale. On the square root scale there is a preponderance of dominance for higher activity in Test 1. The above conclusions are based on the analysis of means: in this case the failure of the analysis of the components of variation merely serves to confirm the presence of interaction.

Jakway and Goy

A fourth set of data susceptible to a complete biometrical analysis has recently become available. It relates to the analysis of sexual behavior in the male and female guinea pig

(Goy & Jakway, 1959; Jakway, 1959). Two highly inbred strains, whose history is documented as far back as 1906, and whose near homozygosity was established as early as 1927 by the method of exchanging tissue grafts (Loeb & Wright, 1927), were crossed, and F_{1s} , F_{2s} , and both backcrosses were bred, reciprocally in each case.

In the study of the inheritance of sexual behavior in *female* guinea pigs, the response to the injection of a controlled amount of female hormones in previously ovariectomized subjects was assessed in terms of four behavioral measures. No details are given of precautions taken to minimize environmental effects. The test technique is described as follows:

The median age at the time of ovariectomy was 3.5 months in each genetic group. The distributions were not skewed. . . . Tests of reproductive performance began one month later on the average. For the first 3 tests, each animal was injected with 100 I.U. of oestradiol benzoate followed 36 hours later with 0.2 I.U. of progesterone. . . . The volume of all injections was constant (0.5 c.c.), and injections were given subcutaneously in the left axilla. Immediately after injection with progesterone, the animals were placed in a standard observation cage (in groups of 6 to 12 individuals) and observed continuously for 14 hours. Each animal was tested once every hour to determine the time of appearance of the lordosis reflex. . . . The first lordosis obtained was regarded as the onset of oestrus, and animals failing to respond on any of the 14 hourly tests were viewed as not in oestrus. For those animals responding on at least one hourly test, oestrus was regarded as terminated when they failed to lordose on two successive hourly tests (Goy & Jakway, 1959, pp. 142-143).

The measures used are detailed in another paper (Goy & Young, 1957) as follows:

(1) Latency of heat is the length of the interval between the injection of progesterone and the elicitation of the first lordosis. (2) Duration of heat is the number of hours lordosis can be elicited. Toward the end of a heat period, lordoses become feeble and difficult to

TABLE 4
THOMPSON AND FULLER'S DATA: MEANS AND THEIR STANDARD ERRORS IN UNITS AS INDICATED AND (n)

Measure	Generation Means					
	P ₁ (C57BR)	P ₂ (A/Jax)	F ₁	F ₂	B ₁	B ₂
Test 1. Number of lines crossed	532.2 ± 23.6 (40)	11.4 ± 5.2 (40)	302.9 ± 16.6 (38)	287.9 ± 16.0 (92)	148.3 ± 14.3 (58)	395.9 ± 18.8 (63)
Test 1. Square root transformation	22.9 ± 0.5	1.9 ± 0.6	17.0 ± 0.6	16.1 ± 0.6	11.0 ± 0.7	18.9 ± 0.7
Test 2. Number of half maze units traversed	248.8 ± 11.1	79.6 ± 6.5	160.6 ± 5.4	164.9 ± 4.4	131.5 ± 3.7	206.2 ± 2.7

elicit, and an operational criterion is necessary to determine when an animal shall be classified as unresponsive. For this purpose, each animal is stroked or fingered five times and if no lordosis is displayed the animal is considered to be out of heat. . . . (3) The duration of maximum lordosis in seconds. . . . The lordosis reflex includes several components, an arching or straightening of the back, elevation of the pudendum, displacement of the rear feet laterally and caudally so that a wide stance is taken, and emission of a low guttural growl. When an estrous female is stroked lightly in a caudo-cephalad direction all components of the lordosis are displayed nearly simultaneously. If the stroking is continued (prolonged stimulation), the full reflex will be maintained for a time which varies with the genetic background and the phase of estrus. If stimulation is continued until voluntary termination is produced, the duration of the reflex can be measured with a stop-watch. The response may be considered terminated when any one of the following signs is evident: (a) a sudden or gradual return of the back and pudendum to a normal position; (b) a sudden return of the feet to the normal position and a loss of the wide stance characteristic of the reflex; (c) kicking with the hind feet; (d) dashing forward; (e) squatting; (f) urinating; and (g) an abrupt termination of the growl accompanied by a soft squeal. A stop-watch is started immediately on display of lordosis and stopped as soon as the complete response is no longer apparent. . . . (4) Male-like mounting behavior. Mounts accomplished by an individual are classified as (a) complete mounts at the posterior end including pelvic thrusts, (b) posterior mounts without pelvic thrusts, and (c) abortive mounts which are not posteriorly oriented, do not involve clasping, and usually are not accompanied by pelvic thrusts. Recorded mounting activity is usually preceded by locomotor activity best described as prowling or standing in one place and treading the floor of the cage with the hind feet. Both treading and prowling, when they precede mounting, are accompanied by the typical low guttural growl or chatter. (5) Per cent of females brought into heat by the hormonal treatment (pp. 342-343).

The means, together with SEs for the first four measures, are given in Goy and Jakway's Table 1,⁸ and are not repeated here.

⁸ The standard errors for the measure "number of mounts per oestrons" were recalculated from the distributions given in

The data . . . were not normally distributed and the variances of the different genetic groups were unequal. Because of these characteristics, conventional parametric analysis was not feasible. Therefore only . . . non-parametric statistics were employed in the analysis (Goy & Jakway, 1959, p. 143).

However, in the case of Measures 2, 3, and 4 above, distributions in the form of proportions are given which has enabled rescaling as necessary.

The methods used with the *male guinea pigs* in rearing, testing, and scoring their sexual behavior are described by Jakway (1959) as follows:

The animals were left with their own dams and siblings until weaning on day 25. They were then placed in individual cages 2 ft. X 2 ft. X 1 ft. with two females of their own age. The caging in such groups assured each male of the contact with other animals which is necessary to bring out the behavioural differences between males of the two inbred strains. On day 73 the female cagemates were removed. Between the ages of 77 and 120 days each animal was observed in seven, approximately weekly, 10-minute tests with oestrus females. The mean score from this number of observations is expressive of the mating performance of a given animal. Elements or measures of sexual behaviour . . . are defined as follows: *Circling* is the term employed when the male circles the female. *Sniffing and nibbling* is recorded each time the nose of the male touches the female other than in the anogenital region. *Nuzzling* is recorded when the nose touches the anogenital region of the female. *Mounting* is scored when the male places both forepaws on the female. *Intromission* is recorded when the penis penetrates the vaginal orifice. This is accompanied by rhythmic pelvic thrusts. *Ejaculation* is accompanied by a convulsive contraction of the haunches and terminates the display of sexual behaviour. A test score is a numerical value reflecting three factors: the interval of ejaculation (latency of ejaculation), the amount of sexually oriented activity, and the maturity level of the behaviour. With the exception of circling which is not scored, each measure is given a numerical value from the

Goy and Jakway's Table 3. Our values are in general agreement with those given in their Table 1, taking account of distortion due to grouping, with the exception of Strain 13. The obviously erroneous value given in the last column of their Table 3 may be responsible for the discrepancy.

lowest for sniffing and nibbling to the highest for ejaculation. The value of each is then multiplied by a factor expressive of latency of ejaculation; the shorter the latency, the higher the factor. Since most tests in which ejaculation occurred were terminated before the end of the tent¹/₂ minute, measurements other than scores will be expressed as rates/15 seconds. Inasmuch as a sexual behaviour score can be attained in several ways, the components were analysed separately for possible patterns of inheritance (p. 151).

The means and standard errors for each component and the composite score are grouped together in Table 5. In each case Jakway gives the percentage distributions which has enabled us to rescale the data as necessary.

The results of the joint scaling tests are summarized in Table 6. Only two measures, *duration of maximum lordosis* in females and *number of ejaculations* in the males show genic interactions on the chosen scales. All the measures show significant heritable variation and only one, *circling* in males shows no significant dominance.

For the four female measures the potency ratio $[h]/[d]$ is approximately half and for three of the measures, *latency of estrus*, *duration of maximum lordosis*, and *frequency of mounting* it is also negative. That is, for these measures the parent with the lower score contains a preponderance of dominant genes. For the male measures the potency ratios are more variable, ranging from non-significant for *circling* to greater than 10 for *number of ejaculations*; in fact for three measures, *nuzzling*, *intromissions*, *ejaculations*, as well as for the composite score, the ratio is greater than one, that is, all these measures show heterosis.

Estimates of the components of the generation means and their standard errors showed that the $[j]$ -type interaction is primarily responsible for the significant deviation for additiv-

ity in *duration of maximum lordosis* with a value of 6.3 ± 2.4 , while it is the $[l]$ -type interaction which gives the same effect in the measure *number of ejaculations* (6.0 ± 2.0). Tests of inhomogeneity of the P_1 , P_2 , and F_1 variances show that three of the female measures, *latency of estrus*, *duration of maximum lordosis* and *frequency of mounting* and two male measures, *circling* and *nuzzling*, exhibit significant genotype-environmental interactions. In all, therefore, the data from six measures would require rescaling to remove either genic or genotype-environmental interaction in order to proceed with the analysis of the second degree statistics.

Unfortunately, the observation on *latency of estrus* in the females are not presented in a manner which allows rescaling. For the other measures both square root and log transformations were made and the latter in all cases removed both sources of non-independence. Thus, for the two measures showing genic interaction the joint scaling tests on the log transformed data gave values of $[d] = 0.138 \pm 0.014$, $[h] = 0.030 \pm 0.028$ for *duration of maximum lordosis*, and $[d] = 0.012 \pm 0.009$, $[h] = 0.126 \pm 0.018$ for *number of ejaculations* both of which now gave satisfactory fits with the observed data. A solution for D , H , and E_1 was therefore attempted from the second degree statistics of all the measures on all scales. These estimates proved to be disappointing. Only two measures, *frequency of mounting* in the females and *intromissions* in the males gave evidence of segregation in the F_2 and backcross generations on any of the three scales employed—in the former case, on two of them. That is, only for these two measures were the magnitudes of V_{D_1} and $V_{D_1} + V_{D_2}$ greater than the estimates of their environmental component (E_1 and $2E_1$, respectively) ob-

TABLE 5
JAKWAY'S DATA: MEANS AND THEIR STANDARD ERRORS IN UNITS AS INDICATED AND (n)

Measure	Generation Means						
	P ₁ (strain 2) (20)	P ₂ (strain 13) (18)	F ₁ (20)	F ₂ (30)	B ₁ (53)	B ₁ Reciprocals	
						P ₁ females (27)	F ₁ females (26)
Circling/15 seconds	0.7 ± 0.05	0.3 ± 0.03		0.4 ± 0.03	0.5 ± 0.03	0.4 ± 0.02	0.6 ± 0.03
Nuzzling/15 seconds	0.7 ± 0.04	0.5 ± 0.04	0.4 ± 0.02	0.6 ± 0.03	0.7 ± 0.02	0.6 ± 0.03	0.7 ± 0.03
Mounting/15 seconds	0.8 ± 0.04	0.3 ± 0.03	0.5 ± 0.02	0.5 ± 0.02	0.6 ± 0.03		0.3 ± 0.02
Intrusions/15 seconds	0.06 ± 0.01	0.04 ± 0.01	0.1 ± 0.01	0.08 ± 0.09	0.07 ± 0.007		0.06 ± 0.005
Ejaculations in 7 tests	1.1 ± 0.3	0.6 ± 0.2	4.9 ± 0.4	1.5 ± 0.3	1.7 ± 0.3	2.5 ± 0.4	0.9 ± 0.20
Sexual behavior score	6.4 ± 0.2	4.6 ± 0.2	7.3 ± 0.3	5.7 ± 0.2	6.2 ± 0.2		4.9 ± 0.1

TABLE 6
JAKWAY AND GOY'S DATA: COMPONENTS OF MEANS AND THEIR STANDARD ERRORS FROM JOINT SCALING TESTS, AND HERITABILITIES

Measure	Components of means		Interaction p from χ^2	Heritability [d]/([d]+e) (%)
	[d]	[h]		
Females				
Latency of estrus (hours)	1.2 \pm 0.1	-0.88 \pm 0.21		87.5
Duration of estrus (hours)	1.6 \pm 0.2	0.99 \pm 0.30		87.2
Duration of maximum lordosis (seconds)	7.3 \pm 0.2	-3.17 \pm 1.16	0.001	89.2
Frequency of mounting	8.7 \pm 0.5	-4.71 \pm 0.88		
Males				
Circling/15 seconds	0.18 \pm 0.06	-0.17 \pm 0.18		83.4
Nuzzling/15 seconds	0.10 \pm 0.02	-0.16 \pm 0.04		77.6
Mounting/15 seconds	0.2 \pm 0.02	-0.12 \pm 0.03		89.4
Intromissions/15 seconds	0.01 \pm 0.005	0.09 \pm 0.01		
Ejaculations in 7 tests	0.34 \pm 0.18	3.7 \pm 0.4	0.001	54.3
Sexual behavior score	0.98 \pm 0.13	1.6 \pm 0.2		83.1

tained from the parental and F_1 variances. The results of these estimations are given in Table 7.

Estimates of heritability are therefore only possible for *frequency of mounting* and *intromissions* and these

lie between 50% and 60% (see Table 7). An estimate of the number of effective factors (K) is also confined to these two measures neither of which give values greater than one, although it is quite clear that more

TABLE 7
JAKWAY AND GOY'S DATA: COMPONENTS OF VARIATION AND HERITABILITY

Measure	Scale	Components			Heritability	
		D	H	E_1	D	$\frac{1}{2}D + \frac{1}{2}H$
					$\frac{D+E_1}{2}$ (%)	$\frac{\frac{1}{2}D + \frac{1}{2}H + E_1}{2}$ (%)
Frequency of mounting (females)	Linear	26.61	13.52	16.63	61.4	50.0
	Log	0.102	0.246	0.088	53.5	56.0
Intromission	Linear	0.0014	0.0016	0.0011	56.0	50.0

than one gene must be involved in the inheritance of the characters which show genic interaction. To obtain some idea of the heritability of the remaining measures a first degree equivalent of one of our estimates has been evaluated, namely, $[d]/([d]+e)$ where e is the mean standard error derived from the sampling errors of the generation means as follows:⁹

$$e = \sqrt{(V_{F_1} + V_{F_2} + V_{F_3} + V_{F_4} + V_{B_1} + V_{B_2})/6}$$

and the results are indicated in Table 6.

The failure of the second degree statistics to show even evidence of segregation for 8 of the 10 measures analyzed requires further comment. A number of contributory factors are present which might reduce or bias our estimates of these statistics as derived from our transformations of the data as published. These include: (a) grouping of the data into as few as five classes, (b) the use of metrics which place over 50% of the individuals into the zero class, and (c) grouping all scores higher than an upper limit set by the higher parent or F_1 into one class which may contain 30% of the individuals of a segregating generation.

The exact consequences of these procedures are difficult to ascertain but it is clear that they lead to scalar problems which have not been resolved by either square root or log transformations, and they could easily reduce the variances of the segregating generations (F_3 , B_1 , and

B_2) to those for the nonsegregating generations (P_1 , P_2 , and F_1).

Our conclusions are therefore necessarily drawn mainly from the analysis of the first degree statistics and are substantially in agreement with those of Goy and Jakway (1959) and Jakway (1959). There are, however, differences in detail. To take one example; they make no allowance for the possibility of genic interaction which is unambiguously present in two of the measures on the original scale. In consequence, they consider that *maximum lordosis* in females is under the control of a single genetic factor without dominance; an interpretation which is difficult to uphold in view of our demonstration of significant [j]-type nonallelic interactions.

Scott

We can now turn to the analysis of less complete sets of data, the most recent among which is that of Scott (1954). This is one small segment of that which has been collected at Bar Harbor on the performance of five thoroughbred strains of dogs. The two strains for which crossbred data have so far been reported are the cocker spaniel and the African basenji or barkless dog. The breeding program used is described as follows:

It was found that all of the breeds showed a great deal of variability, a large part of which appeared to be hereditary since offspring of different matings gave different results. In order to reduce this variability somewhat, the animals chosen from the parent strains for the crossbreeding experiment were descended from one brother \times sister mating in the basenji's and from two matings of a single male with his sister and mother in the case of the cocker spaniels. No selection of these individuals was used except that the original pairs were vigorous and healthy animals. As it turned out later these did not necessarily illustrate the extremes of either breed in all characteristics. Reciprocal crosses were made between these two groups of siblings and an effort was made to obtain at least four dif-

* This formula is only applicable in this case because of the demonstrated absence of heritable variation in the segregating F_2 and backcross generations. These may therefore be treated as equivalent to the parental and F_1 generations in displaying only environmentally induced variation, although, of course, they would normally be excluded from this type of estimate.

ferent pairs in each case, giving two F_1 populations. F_1 males were backcrossed to the mothers so that backcross and F_1 animals raised by the same mothers could be paired. Finally, F_2 populations are being obtained from both crosses (Scott 1954, p. 745).

The subjects were reared in a carefully standardized manner as part of the program, and subjected to a battery of psychological and physiological tests at various predetermined stages in their life history. Scott (1954) gives the results of analyses involving nine different measures, but only in one case is the grouped data for the distributions of the scores given, thus enabling calculations of the approximate means and variances for the various generations. These were the scores derived from a barrier test given the pups at the age of 6 weeks on 2 days. The task is to seek the way round a barrier to reach the experimenter and food.

The F_2 data have not yet been published, so that the results to be found in Table 8 relate to the parental, F_1 and backcross lines only. With both backcross means higher than either of the parental or F_1 means, no scale is possible on which genic interaction is absent. It is not surprising, therefore, to find significant deviations from additivity of gene action on the two available scaling tests, the A and B tests ($p=0.01$ and 0.05 , respectively). Nor

is this situation improved by a square root or a log transformation.

In the absence of the F_2 generation mean we can only estimate [d] and [j] among the components of the means. In such a situation, we can, however, obtain estimates of various compounds of the remaining components. Thus,

$$[i] - [j] = \bar{F}_1 - \frac{1}{2}(\bar{P}_1 + \bar{P}_2)$$

$$[i] + [1] = \bar{P}_1 + \bar{P}_2 + 2\bar{F}_1 - 2\bar{B}_1 - 2\bar{B}_2$$

$$[h] + [1] = \frac{1}{2}\bar{P}_1 + \frac{1}{2}\bar{P}_2 + 3\bar{F}_1 - 2\bar{B}_1 - 2\bar{B}_2$$

Scott's data give the following values for these components on the linear scale:

$$[d] = 1.7 \pm 3.9$$

$$[j] = -4.9 \pm 8.4$$

$$[h] - [i] = -1.3 \pm 1.9$$

$$[i] + [1] = -30.5 \pm 8.6$$

$$[h] + [1] = -31.8 \pm 8.6$$

Only [i] + [1] and [h] + [1] are significant; therefore the only certain feature of these data is the presence of nonallelic interactions.

From the second degree statistics we can estimate only $D+H$, $\Sigma(dh)$, and E_1 , but we cannot place much reliance on their values in the presence of both unscalable nonallelic interactions and slight genotype-environmental interaction ($p=0.05$).

TABLE 8

SCOTT'S DATA: MEANS AND THEIR STANDARD ERRORS IN NUMBER OF ERRORS AND (n)

	Generation Means				
	P_1 Basenji	P_2 Cocker Spaniel	F_1	B_1	B_2
Actual Parents	2.5 ± 1.7 (16)	10.8 ± 2.3 (26)	5.4 ± 1.0 (41)	12.8 ± 2.7 (27)	14.5 ± 2.9 (23)
Total Population	3.2 ± 1.0 (39)	12.5 ± 1.7 (49)			

A square root transformation removes the latter without, as has been noted, removing the genic interaction. On this scale:

$$D+H=0.329$$

$$\Sigma(dh)=0.008$$

and

$$E_1=0.084$$

Although we cannot separate D and H , we can estimate the likely order of heritability by putting $D=H$ and $H=0$ in turn which lead to estimates of 66% and 80%, respectively. The value of $\Sigma(dh)$ provides no evidence of dominance and if the genes show dominance the dominant alleles must be equally distributed between the two parents. With a nonsignificant estimate of $[d]$, on all scales tried, no estimate of the number of effective factors has been attempted. The presence of genic interactions, however, shows that a number of genes are involved.

Tryon

A further set of data are provided by Tryon (1929, 1940, 1942), whose study of selective breeding for "maze-brightness" and "maze-dullness" in rats is probably the best known in the whole of psychogenetics. He selected through 22 generations of brother \times sister mating for high and low error scores in a 17-unit automatic maze in which the rats were trained for 19 trials to run to a food reward. He claims that

Rigorous environmental controls were effected (1) by instituting standard procedure of ani-

mal care and of breeding, (2) by using an automatic mechanical device for delivering the animals into the maze without handling, and (3) employing an electric recorder for the scoring of each rat's maze run (Tryon, 1940, p. 112).

Elsewhere Tryon (1931) details the husbandry and comments as follows:

Very special efforts were made to keep ambient influences the same for all the cages in which these animals lived before they learned the maze (Tryon, 1929). Each animal lived with its siblings until shortly after weaning time (30 days), when it was numbered by punching its ears. Then it was placed with 4 animals from other litters in a cage in which it lived until it ran the maze. Each living cage possessed an ever present supply of food and water. All cages were cleaned at the same time and in the same manner. Even so, it would be naive to suppose that life within a cage was identical for all animals. Any rat experimenter knows that social life within a cage is variable and complex. But it would not seem likely that the difference in experience of different rats in the same cage would to any significant degree cause differences in the later learning of the maze under the remote solitary experimental conditions (p. 316).

A cross was made between the two strains developed, and the F_1 and F_2 generations bred, though it is not clear at what stage in the selection experiment this was done. Tryon (1940) only gives the results in the form of histograms showing the percentage of subjects having a particular error score on his "normalized" scale but from these it has been possible to make approximate reconstructions of the original distributions. The results obtained in this way are given in Table 9.

In the absence of the backcrosses we can apply only the C scaling test

TABLE 9

TRYON'S DATA: MEANS AND THEIR STANDARD ERRORS IN NUMBER OF ERRORS AND (n)

Generation Means			
P_1 (bright)	P_2 (dull)	F_1	F_2
25.9 ± 0.9 (85)	142.9 ± 3.7 (53)	63.1 ± 3.5 (133)	71.2 ± 2.9 (202)

and this suggests that on the chosen scale genic interactions are absent. We can therefore estimate [d] and [h] in the manner discussed in the next section and which give values of 117.0 ± 3.8 and -21.3 ± 4.0 , respectively. Thus we have a large significant additive effect and a small but significant dominance contribution.

Unfortunately, there is significant genotype-environmental interaction ($p < 0.01$), which, in the absence of significance in the C scaling test, we will not attempt to scale out.¹⁰ Our analysis of the second degree statistics is thus prospectively biased. In any case only estimates of $\frac{1}{2}D + \frac{1}{4}H = 646.63$, and $E_1 = 1006.01$ can be obtained, giving the percentage of heritable variation in the F_2 population as 39. If we assume $D = H$ or $H = 0$, we can (a) estimate the number of effective factors as 14.1 and 10.6, and (b) obtain values for our other index of heritability of 49% and 56%, for these two situations, respectively.

Thus the pattern of rat behavior investigated by Tryon is controlled by many genes which are additive in their effect but interact with the environment. They show dominance and there is a preponderance of dominant genes for a low score. The heritabilities for this character are about average.

Vicari

A further set of data which omit the backcrosses are provided by the investigations of Vicari (1929). These are the earliest psychogenetical ex-

periments which yield any data amenable to analysis by the methods proposed here. She used four strains of mice which were "closely inbred" and derived F_1 and F_2 generations from them. One of these strains, the Japanese waltzer, was regarded at the time as being of a different species from the other three (*Mus musculus*), and the offspring of the cross involving it consequently hybrids. It has since been shown (see Grüneberg, 1952), however, that the Japanese waltzer is a subspecies of *Mus musculus*, and furthermore, that it differs from the normal in a genetically complex manner, that is, the waltzer condition is not due to a single gene difference. Vicari's measures were derived from a simple, two-choice maze in which the subjects ran to a food reward. No details of the deprivation schedule for motivating the animals to run this maze are given, and it is clear from the running times reported that the apparatus itself was ill-designed for the purpose of obtaining efficient learning. Despite these difficulties, however, Vicari reports the results after 14 trials for a substantial number of subjects from the four parental generations and the three F_1 s and F_2 s bred from them. We can, of course, only apply the C scaling test, the results of which are given along with the generation means in Table 10. Where there is a significant deviation on the scaling test only estimates of the following compounds and their standard errors can be obtained,

$$[d] - \frac{1}{2}[j] = \frac{1}{2}(\bar{P}_1 - \bar{P}_2)$$

and

$$[h] - [i] = \bar{F}_1 - \frac{1}{2}(\bar{P}_1 + \bar{P}_2)$$

as before, and from Scaling Test C itself:

$$2[i] + [l] = \bar{P}_1 + \bar{P}_2 + 2\bar{F}_1 - 4\bar{F}_2$$

¹⁰ Since, for most of the data previously considered in this paper, rescaling has been necessary because of a failure of both scaling criteria, no attempt has been made to scale out the significant genotype-environmental interaction in this case on the grounds that improvement in respect of the latter disturbance might be at the expense of the satisfactory outcome of the test for additivity.

TABLE 10
VICARI'S DATA: MEANS AND THEIR STANDARD ERRORS OF MEASURES SHOWN AND (*n*)

Measure	Quantity	Generation Means									
		Japanese Wistar Ky (28)	Cross $P_1 \times P_2$		Albino P_3 (78)	Cross $P_1 \times P_2$		Dilute Brown P_4 (27)	Cross $P_1 \times P_2$		Brown (Abnormal eye) P_5 (26)
Number of trials to first errorless run	Mean	8.8 ± 0.8	F_1 (119)	F_2 (46)	8.7 ± 0.4	F_1 (31)	F_2 (45)	6.9 ± 0.7	F_1 (14)	F_2 (62)	10.1 ± 0.8
	C Scaling Test* Genotype-environmental interaction (d) - (f) (b) - (e)		9.2 ± 0.4 -6.1 ± 1.6**	2.0* 0.05 ± 0.45 0.45 ± 0.6		7.2 ± 0.7 8.1 ± 2.5**	5.5 ± 0.9		6.1 ± 0.6 -3.3 ± 1.4**	8.1 ± 0.5	
Largest number of consecutive er- rorless runs	Mean	0.9 ± 0.2	2.0 ± 0.2 2.5 ± 1.4	117.7** 0.25 ± 0.7 0.85 ± 0.3	1.4 ± 1.3	1.4 ± 0.2 -1.3 ± 1.4	1.8 ± 0.1	1.5 ± 0.1	1.8 ± 0.2 -1.5 ± 1.3	1.9 ± 0.2	1.1 ± 0.2
	C Scaling Test Genotype-environmental interaction (d) - (f) (b) - (e)										
Total number of errorless runs	Mean	1.5 ± 0.3	2.7 ± 0.3 2.9 ± 0.8**	3.7** 0.2 ± 0.15 1.0 ± 0.4	1.9 ± 0.2	2.6 ± 0.4 -2.5 ± 0.5**	3.1 ± 0.3	2.8 ± 0.3	3.2 ± 0.6 0 ± 1.5	2.7 ± 0.3	1.7 ± 0.4
	C Scaling Test Genotype-environmental interaction (d) - (f) (b) - (e)										
Mean time per run in seconds	Mean	73.1 ± 8.6	41.9 ± 3.7 -75.9 ± 20.9**	1.4 11.5 ± 4.8 -19.7 ± 5.9	50.1 ± 4.4	15.8 ± 1.6 16.9 ± 7.3*	19.2** 18.8 ± 2.4 -15.5 ± 2.9	12.5 ± 1.7	13.9 ± 2.1 15.5 ± 5.8*	12.9 ± 2.1	27.0 ± 2.9
	C Scaling Test Genotype-environmental interaction (d) - (f) (b) - (e)										

* 2(f) + (1) (see text).

* Significant at the 5% level.

** Significant at the 1% level.

TABLE 11

VICARI'S DATA FOR JAPANESE WALTZER X ALBINO CROSS: MEANS AND THEIR STANDARD ERRORS IN MEAN RUNNING TIME AND (s)

Trial	Generation Means				C Scaling Test	Genotype-environmental interaction
	Japanese Waltzer	Cross F ₁ ×P ₂		Albino		
		P ₁	F ₁			
1	113.7 ± 12.0 (80)	74.2 ± 7.4 (151)	86.8 ± 11.7 (61)	92.3 ± 11.0 (80)	7.2 ± 32.2	1.5**
4	73.4 ± 8.6 (68)	67.3 ± 7.2 (149)	70.6 ± 11.4 (61)	60.2 ± 8.7 (79)	-12.0 ± 29.7	1.5**
8	52.3 ± 10.0 (45)	47.0 ± 5.9 (128)	60.3 ± 10.0 (53)	47.1 ± 5.7 (78)	-47.7 ± 26.0	1.3*
14	83.0 ± 16.8 (28)	21.1 ± 2.6 (119)	51.7 ± 11.6 (49)	29.6 ± 4.6 (78)	-52.2 ± 29.4	9.7**

* Significant at the 5% level.

** Significant beyond the 1% level.

But direct estimates of [d] and [h] can be obtained where there is no significant deviation on the scaling test by assuming [i] and [j]=0, as in the case of Tryon's data above.

Of the 12 sets of data, i.e., four measures recorded for each of the three crosses, eight show significant genic-interaction and eight, genotype-environmental interaction on the chosen scales. Unfortunately we cannot rescale the data because Vicari does not give the individual scores on which the means and variances are based. She does, however, present the distributions for one measure, mean running time, for different stages in the experiment, the first, fourth, eighth and fourteenth trials. We shall use the cross $P_1 \times P_2$ to illustrate the further analyses, the relevant generation means appearing in Table 11. The decrease in the number of subjects in successive trials which is observed is attributed by Vicari to death, escapes, failures to run, etc. An analysis of variance of the 4×4 table, that is, four generation means in each of four trials, shows a highly significant difference between trials ($p < 0.01$) and a significant difference between generations ($p = 0.05-0.01$) when compared with the interaction mean square for generations \times trials, which has the same

order of magnitude as the sampling variance of the generation means. The significant difference between generations is expected if the character is inherited and the significant difference between trials, the running time falling steadily as the number of trials increases, reflects a strong training component which is presumably non-heritable. Since the C scaling test when applied to these data detected no significant deviations due to genic-interaction (Table 11), we can estimate [d] and [h] directly from the generation means on the original scale.¹¹ For the first trial [d]=10.7 \pm 8.2 and [h]= -28.8 \pm 17.8 neither of which are significant, that is, there is no significant heritable component of the generation means. The fully trained performance in the fourteenth trial gives [d]=26.7 \pm 8.7 and [h]= -35.2 \pm 9.1. The components are similar in absolute and relative magnitudes in these two extreme cases but their significance is higher in the last trial. This greater heritability in the last trial is supported by the second degree statistics. Thus in the first trial the percentage of heritable variation in the F_2 population

¹¹ In ignoring the significant genotype-environmental interactions shown in Table 11, we are following the argument of Footnote 9.

is not significantly different from zero, while in the fourteenth trial it is 51%. Furthermore the fourth and eighth trials in this case give intermediate values of 21% and 29%, respectively. Hence heritability increases almost linearly with the number of trials, the final performance being more heritable than the initial performance.

Before attempting an interpretation of this finding two points must be borne in mind. Firstly, the experimenter's skill may have improved in successive trials, thus reducing the nonheritable component of variation, though this is unlikely since the generations were probably not all tested at the same time, and, secondly, almost 30% of the animals scored in the first trial were missing in the fourteenth for a number of reasons. If the missing third were not a random sample of the original subjects a progressive bias could have been introduced. It is not possible to ascertain from the data available whether or not one or both of these factors is making a contribution to the observed trend. It is interesting, however, to note that the total variation in the F_2 population remained constant from the first to the last trial and the increased heritability results from a drop in the percentage due to nonheritable agencies from 100 to 49. It seems likely, therefore, that what we have detected is in fact a real effect and that it represents a progressive release of the performance from the effect of environmental stimuli irrelevant to it.

This effect of a progressive increase in the heritability of performance indicated in Vicari's data would seem, if confirmed, to have interesting implications. It might be related to the change over from general to specific factors known to occur in the acquisition of skills (Fleishman, 1957;

see also Wherry, 1939), and might also indicate a method for assessing the relative importance of environmental variation in learning tasks. The effect of the *same* environmental stimuli at different stages in a given task might be studied, as in Vicari's situation, or at the same stage in different tasks, as well as that of *different* stimuli in either of such arrangements. As was noted earlier, the sort of problems which the inclusion of environmental variation introduces into biometrical analyses has been discussed and possible solutions indicated (Jones & Mather, 1958; Mather & Jones, 1958; van der Veen, 1959). While it is beyond the scope of this paper to enter into this matter in detail, it may be said that there is evidence for a genetical component in the determination of the *variability* of performance in such different environments, as opposed to the control of its actual expression in any single one of them which is what we have been dealing with so far. This variability is also susceptible to analysis by biometrical methods (Jinks & Mather, 1955). The analysis from this point of view of the only suitable behavioral data at present known to us (Broadhurst, 1960) is not yet complete.

Thus, while the nature of Vicari's data makes our conclusions tentative, the analyses discussed here have shown their advantages if only in indicating the complexity of the inheritance of the behavior patterns under investigation.

DISCUSSION

We have presented the results of our reanalyses of all the data available to us in the field of psychogenetics with little reference to the outcome of analyses of other kinds performed, in some cases, by the authors concerned. The methods

used have usually been those of classical Mendelian genetics, which, though basic to biometrical genetics, cannot satisfactorily be applied in their simpler forms to the analysis of continuously variable characteristics. The clarity associated with the Mendelian analysis of discontinuous variation and attributable thereby to the effects of two or three major genes is not to be expected from biometrical methods, a major assumption of which is that the continuous phenotypic variation observed is the product of multiple genetical and environmental causes, largely unspecified in detail. Sometimes the argument from Mendelian analysis has been by analogy, which is not always illuminating and may in certain instances be downright misleading. For example, the resemblance of the F_1 to one or other of the parental lines does not necessarily mean the same thing in biometrical genetics as it does in the simpler cases encountered in Mendelian analysis. According to the polygenic hypothesis, the departure of the F_1 mean from the mid-parental value depends both on the balance in the parental lines of the dominant and recessive polygenes and upon the respective *direction* of their cumulative effects. Dominants may be increasers or decreasers, that is, having a positive or a negative phenotypic effect, respectively, as expressed on the scale used, or there may be a balance between the two. Thus the F_1 mean value may be close to the upper parental mean on a scale because of a preponderance in the parents of dominants with positive effect in terms of the metric, while an F_1 close to the lower parental value will result from a preponderance of dominant decreasers. Intermediate values may express different degrees of balance operating. Failure to recognize this difference between

potence, measured in terms of $[h]$, and dominance, measured as \sqrt{H} , will in general lead to spurious disagreements between the level of dominance in the F_1 s and that in the segregating F_2 and backcross generations (e.g., Jakway, 1959, p. 155). It is equally misleading to regard low potence, that is an F_1 close to the mean parental value, as a diagnostic characteristic of multifactorial inheritance (see Hall, 1951, p. 321). Genetical methods and analyses more complex than those that have been considered here, such as the analysis of double back-crossing (Mather, 1949), or of diallel crosses (Broadhurst, 1959, 1960; Hayman, 1954; Jinks, 1954), are needed to enable more precise estimates to be made of the different factors operating. In this connexion it should be noted that the analysis in turn of single crosses from two parental strains such as we have been dealing with here is a relatively inefficient and laborious method of making a biometrical analysis of quantitative data. Techniques involving crossing several pure strains at once are superior, and of these the diallel cross method in which a number of strains are intercrossed reciprocally in all possible combinations and analyzed, together with the parental lines, in a single diallel table is probably the best. The merits of this approach at the present time in psychogenetics have been argued elsewhere (Broadhurst, 1960); they are, briefly, that the analysis can proceed at the F_1 stage, without the necessity of breeding further generations, and that the method thereby provides a quick survey in several parental strains at once of the genetical determinants of the character investigated. The former may be of particular utility in interspecies crosses, where hybrids are sometimes sterile, so precluding the possibility

of breeding F_{28} s, etc. Intensive study of the gene differences in pairs of strains selected for further analysis in this way can then follow.

One source of satisfaction from the present work has been the consistently high heritabilities obtained from our analyses; and by choosing a scale, wherever possible, on which interactions between genes, and between genes and environment are absent, these estimates are somewhat more reliable as well as higher than those obtained by the authors of the papers under review. Any inefficiencies in the experimental design—e.g., the imperfect control of environmental variation, or the unreliability of the measures used—or in the statistical analysis used will reduce the values obtained below their true values. While it has been beyond our power to influence the first source of inefficiency in these data, our analyses have improved the position regarding the second. And the collaboration between psychologist and geneticist will ensure that future researches are designed with due care and attention to the importance of the relevant genetical principles (see Broadhurst, 1960).

The choice of scale is an important problem in biometrical analysis (see Mather, 1949) and, as we have seen, the need for rescaling, resulting from interactions between the genes and between the genes and the environment, arises in the inheritance of some 50% and 70%, respectively, of the measures in the examples reviewed here. In no case was the presence of genic interaction demonstrated as a significant factor prior to the analyses undertaken for this review and in only one case were steps taken to eliminate the gene-environment type of interaction by a scalar change (Thompson & Fuller, see Footnote 3). And yet our analyses

show that in almost all cases a simple log transformation is sufficient to eliminate both causes of interaction and hence provide a scale on which unambiguous interpretations of dominance and potency effects can be made.

The gene-environment interaction detected by the inhomogeneity of the variances of the parents and F_1 s has two main causes. The major cause (some 60% of measures) is a correlation between mean and variance in the nonsegregating generations. A more interesting cause, however, in the remaining examples is the lower variance of the F_1 individuals compared with those of the parental generations; an effect which is independent of their means. This phenomenon, which is common to the inheritance of all types of characters and occurs equally among animals and plants, has received considerable attention of late (see Jinks & Mather, 1955; Lerner, 1954; Mather, 1953, for reviews). Extensive discussion of this point is, however, beyond the scope of the present review.

Some 70% of the genic interactions in these analyses are due to the [j]-type interactions. This has two implications. Firstly, there must be interactions between additive and dominance effects and, secondly, the interacting genes must be associated in the parental lines, the majority of increasing interacting genes being present in one parent and the decrease in the other. This is not an unexpected result when one considers that most of the experiments reviewed here have employed parental lines which were chosen because they represented the extreme phenotypes immediately available or obtainable as a result of prolonged selection.

This policy could explain two further features of these examples,

namely, the rarity of heterosis and the often satisfactory estimates of the number of genetical factors attained by a method which, for reasons mentioned earlier and discussed more fully by Mather (1949), have often failed to give sensible values in other work. Without going into details it is clear that if the better parent in a cross already contains the majority of the available increasing genes it is unlikely to give rise to a superior F_1 irrespective of the dominance or interactive properties of the genes. Similarly, our estimate of the number of genetical factors assumes that the genes are associated in the parental lines. Failure of this assumption leads to underestimation. Our rather satisfactory estimates could, therefore, be a further indication that the genes are so distributed in the parental lines.

It is hoped that the reanalyses reported here serve as another example of the application of biometrical methods to psychological data in addition to those already available (Broadhurst, 1959, 1960). Considering the unsatisfactory nature of much of the data at our disposal, it is felt that the outcome, in terms of the ease of the analyses, especially with regard to the search for suitable

scales, and the consistency of the results obtained, has been favorable. It is not yet possible to pronounce on the general efficiency in this field of the methods advocated by us. Further proof of their suitability will only come when they have been applied more widely to data gathered from suitably designed experiments, perhaps along the lines indicated, so that replication of the genetical picture becomes possible, thus enabling some specification to be made of the generality of the determinants of a particular behavioral characteristic.

SUMMARY

The techniques which can be used in the analysis of quantitative data by the methods of biometrical genetics were outlined and the importance of achieving a suitable scale noted. The body of the paper consists of descriptions of experiments in psychogenetics which lend themselves to this type of analysis, and a presentation of the results of our reanalyses of the data they provide in terms of additive, dominance and interaction components of variation. We conclude that, despite the unsuitable nature of some of the available data, the outcome indicates the utility of the biometrical approach.

REFERENCES

- BROADHURST, P. L. Application of biometrical genetics to behaviour in rats. *Nature, Lond.*, 1959, 184, 1517-1518.
- BROADHURST, P. L. Experiments in psychogenetics: Applications of biometrical genetics to the inheritance of behaviour. In H. J. Eysenck, (Ed.), *Experiments in personality*: Vol. I. *Psychogenetics and psychopharmacology*. London: Routledge & Kegan Paul, 1960. Pp. 1-102.
- BROADHURST, P. L. Analysis of maternal effects in the inheritance of behavior. *Anim. Behav.*, 1961, in press.
- BRODY, ELISABETH G. Genetic basis of spontaneous activity in the albino rat. *Comp. Psychol. Monogr.*, 1942, 17, No. 5, 1-24.
- CAVALLI, L. L. An analysis of linkage in quantitative inheritance. In E. C. R. Reeve & C. H. Waddington (Eds.), *Quantitative inheritance*. London: Her Majesty's Stationery Office, 1952. Pp. 135-144.
- DAWSON, W. M. Inheritance of wildness and tameness in mice. *Genetics*, 1932, 17, 296-326.
- FLEISHMAN, E. A. A comparative study of aptitude patterns in unskilled and skilled psychomotor performances. *J. appl. Psychol.*, 1957, 41, 263-272.
- FULLER, J. L., & THOMPSON, W. R. *Behavior genetics*. New York: Wiley, 1960.
- GOV, R. W., & JAKWAY, JACQUELINE S. The inheritance of patterns of sexual behaviour in female guinea pigs. *Anim. Behav.*, 1959, 7, 142-149.

- GOY, R. W., & YOUNG, W. C. Strain differences in the behavioral responses of female guinea pigs to alpha-estradiol benzoate and progesterone. *Behaviour*, 1957, 10, 340-354.
- GRÜNEBERG, H. *The genetics of the mouse*. (2nd ed.) The Hague: Martinus Nijhoff, 1952.
- HALL, C. S. The genetics of behavior. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 304-329.
- HAYMAN, B. I. The theory and analysis of diallel crosses. *Genetics*, 1954, 39, 789-809.
- HAYMAN, B. I., & MATHER, K. The description of genic interactions in continuous variation. *Biometrics*, 1955, 11, 69-82.
- JAKWAY, JACQUELINE S. The inheritance of patterns of mating behaviour in the male guinea pig. *Anim. Behav.*, 1959, 7, 150-162.
- JINKS, J. L. The analysis of continuous variation in a diallel cross of *Nicotiana rustica* varieties: I. The analysis of F_1 data. *Genetics*, 1954, 39, 767-788.
- JINKS, J. L., & JONES, R. M. Estimation of the components of heterosis. *Genetics*, 1958, 43, 223-234.
- JINKS, J. L., & MATHER, K. Stability in development of heterozygotes and homozygotes. *Proc. Roy. Soc. Lond., Ser. B.*, 1955, 143, 561-578.
- JONES, R. M., & MATHER, K. Interaction of genotype and environment in continuous variation: II. Analysis. *Biometrics*, 1958, 14, 489-498.
- LEARNER, I. M. *Genetic homeostasis*. Edinburgh: Oliver & Boyd, 1954.
- LOEB, L., & WRIGHT, S. Transplantation and individuality differentials in inbred families of guinea pigs. *Amer. J. Pathol.*, 1927, 3, 251-283.
- MATHER, K. *Biometrical genetics: The study of continuous variation*. London: Methuen, 1949.
- MATHER, K. Genetical control of stability in development. *Heredity*, 1953, 7, 297-336.
- MATHER, K., & JONES, R. M. Interaction of genotype and environment in continuous variation: I. Descriptions. *Biometrics*, 1958, 14, 343-359.
- PEARSON, E. S., & HARTLEY, H. O. *Biometrika tables for statisticians*. Vol. I. (2nd ed.) Cambridge: Univer. Press, 1958.
- RUNDQUIST, E. A. Inheritance of spontaneous activity in rats. *J. comp. Psychol.*, 1933, 16, 415-438.
- SCOTT, J. P. The effects of selection and domestication upon the behavior of the dog. *J. Nat. Cancer Res. Inst.*, 1954, 15, 739-758.
- SNEDECOR, G. W. *Statistical methods applied to experiments in agriculture and biology*. (5th ed.) Ames: Iowa State Coll. Press, 1956.
- THOMPSON, W. R. The inheritance of behaviour: Behavioural differences in fifteen mouse strains. *Canad. J. Psychol.*, 1953, 7, 145-155.
- THOMPSON, W. R. The inheritance of behavior: Activity differences in five inbred mouse strains. *J. Hered.*, 1956, 47, 147-148.
- TRYON, R. C. Genetics of learning ability in rats: Preliminary report. *U. Calif. Publ. Psychol.*, 1929, 4, 71-89.
- TRYON, R. C. Studies in individual differences in maze ability: IV. The constancy of individual differences: Correlation between learning and relearning. *J. comp. Psychol.*, 1931, 12, 303-345.
- TRYON, R. C. Genetic differences in maze-learning ability in rats. *Yearb. Nat. Soc. Stud. Educ.*, 1940, 39(1), 111-119.
- TRYON, R. C. Individual differences. In F. A. Moss (Ed.), *Comparative psychology*. (2nd ed.) New York: Prentice-Hall, 1942. Pp. 330-365.
- VAN DER VEEN, J. H. The 2×2 genotype-environment table. *Heredity*, 1959, 13, 123-126.
- VICARI, EMELIA M. Mode of inheritance of reaction time and degree of learning in mice. *J. exp. Zool.*, 1929, 54, 31-88.
- WHERRY, R. J. Factorial analysis of learning dynamics in animals. *J. comp. Psychol.*, 1939, 28, 263-272.
- WIGAN, L. G. Balance and potency in natural populations. *J. Genet.*, 1944, 46, 150-160.
- WRIGHT, S. The genetics of quantitative variability. In E. C. R. Reeve & C. H. Waddington (Eds.), *Quantitative inheritance*. London: Her Majesty's Stationery Office, 1952. Pp. 5-41.

(Received June 9, 1960)

TEACHING MACHINES:

A REVIEW¹

CHARLES S. MORRILL²

MITRE Corporation

The same forces which have characterized the evolution of general educational practices are inherent to the history of the new science of automated teaching. As a result of the expansion and multiplying complexities of political, economic, and social interests, there developed an ever increasing need for the rapid education of large numbers of people. New educational objectives demanded new methods of instruction, and the history of education is marked by many diverse attempts at establishing more efficient teaching procedures. Once again teaching methods must be re-evaluated. Rigid adherence to the principle of personal teacher-student relationships no longer seems feasible—an instructional system more appropriate for present-day needs must be established. It is probable that the use of automated teaching devices can fill this need in the method of education: As Corrigan (1959) has suggested:

the automated teaching method has grown out of a pressing need. This need has been created by a twofold technical training problem. As advances in science and technology have been made, there has been an ever increasing demand for well-trained instructors; at the same time the availability of these trained persons has been diminishing. This situation is aggravated further by the increased scope and complexity of subjects, and the ever increasing ratio between number of instructors and students (p. 24).

¹ The research reported in this document was supported by the Department of the Air Force under Air Force Contract AF-33-(600)39852.

² The author wishes to acknowledge the valuable editorial assistance of Sylvia Pilsucki.

CURRENT TRENDS IN AUTOMATED TEACHING MACHINES

Current interest in the area of automated teaching machines is well illustrated by the simple index of frequency-per-year of published teaching machine articles. Fry, Bryan, and Rigney (1960) report that for the years prior to 1948 there are only 6 references, whereas through 1959 there were more than 50 reports published.

The grandfather of automated teaching machines is Sydney L. Pressey (1926, 1927), who designed machines for automated teaching during the mid-1920s. His first device was exhibited and described at the American Psychological Association (APA) meetings in 1924; an improved device was exhibited in 1925 at the APA meetings. Both forms of the apparatus automatically performed simultaneous administration and scoring of a test and taught informational and drill material. Pressey's device, about the size of a portable typewriter, presented material to the subject via a small window. Four keys were located alongside the apparatus. If the student activated a key corresponding to the correct answer, the machine advanced to the next item. If his response was incorrect, the machine scored an error and did not advance to the next item until the correct answer was chosen. The capacity of the drum was 30 two-line typewritten items; the paper on which the questions appeared was carried as in a typewriter.

In 1927, Pressey summarized his efforts as follows:

The paper reports an effort to develop an apparatus for teaching drill material which (a) should keep each question or problem before the learner until he finds the correct answer, (b) should inform him at once regarding the correctness of each response he makes, (c) should continue to put the subject through the series of questions until the entire lesson has been learned, but (d) should eliminate each question from consideration as the correct answer for it has been mastered (p. 552).

In 1930, Peterson devised a self-scoring, immediate feedback device. The Chemo Card, as this device was later called, utilized the technique of multiple choice. A special ink was used by the student in marking his answer. The mark appeared red if the answer was incorrect; a dark color resulted if the answer was correct. Although Pressey's notions and the Chemo Card might have stimulated an interest in automated teaching techniques in the twenties, educators and researchers obviously were not at that time ready for this advanced concept of teaching. Automated teaching did not take hold.

In 1932, Pressey published an article describing a kind of answer sheet which could be scored by an automatic scoring device. This apparatus recorded errors by item, and thus provided the instructor with clues as to what questions needed further instruction. In 1934, Little experimented with this device as well as with the device originated by Pressey in 1926. His results favored the use of automated devices in contrast to regular classroom techniques.

The next appearance of automated teaching literature came a considerable number of years later. During World War II, the Automatic Rater was used by the Navy for training. This device projected a question on

a small screen; the subject's response consisted of pushing one of five buttons.

In 1950, Pressey described a new automated device called the Punchboard. Multiple-choice questions were presented to the student. The key answer sheet inside the Punchboard contained holes opposite the correct answers only. If the answer was correct, the student's pencil penetrated deeply; if incorrect, the pencil did not penetrate the paper significantly. Angell and Troyer in 1948 and Angell in 1949 reported the results of using the Punchboard. Both studies suggested the superiority of this method over traditional classroom procedures.

In 1954, Skinner published "The Science of Learning and the Art of Teaching," which provided the basis for the development of his teaching machines. In this article, he stressed the importance of reinforcement in teaching and suggested teaching machines as a method of providing this needed reinforcement for the learner.

Reports concerning the Subject-Matter Trainer began to appear in 1955 (Besnard, Briggs, Mursch, & Walker, 1955; Besnard, Briggs, & Walker, 1955). This electromechanical device is a large multiple-choice machine used essentially for training and testing in the identification of components and in general verbal subject matter. Extensive research has been done with this device because of its considerable flexibility, i.e., it allows several modes of operation for self-instruction: variety of programed subject matter, drop-out feature after items have been mastered, etc.

The Pull-Tab, used experimentally by Bryan and Rigney in 1956, was a device in which the subject received not only a "right" or "wrong" indica-

tion after his choice but also a somewhat detailed explanation of "why" a response was incorrect. In 1949, Briggs had found in experimenting with the Punchboard that learning is significantly enhanced by immediate knowledge of results. Bryan and Rigney's data illustrated that the combination of immediate knowledge of results plus explanation, if the student is in error, produced significantly higher scores on a criterion test than if no explanation had been given. The importance of this research from a historical point of view is that it investigated immediate knowledge of results as a factor existing on a continuum with varying degrees of effect. Up to this point any comparison involving the effectiveness of teaching machines had been one between classroom instruction and the "new" machine under consideration. In Briggs' and in Bryan and Rigney's research, however, we see the beginning of a concern, to become greater in the next few years, with the possible effects of specific variables and their interactions on learning.

The years 1957-58 mark the beginning of the period in which resurgent interest in teaching machines was initiated. Ramo's arguments (1957) reopened the consideration of automated techniques for classroom use. His article served as one of the more forceful attempts to alert educators to the needs and requirements for automated techniques in education. Skinner's continued interest (1958) served as the major catalyst in this area. In his article, he reviewed earlier attempts to stimulate interest in teaching machines and further explained that the learning process was now better understood and that this increased sophistication would be reflected in teaching machine tech-

nology. Skinner suggested that the most appropriate teaching machine would be that which permits the student to *compose* his response rather than to select it from a set of alternatives. On the basis of this philosophy and in conjunction with other principles of learning theory to which Skinner adheres, he designed a teaching machine with the following characteristics. The questions, printed on a disk, are presented to the student through a window. The student's response is written on a paper tape, which is advanced under a transparent cover when the student lifts a lever. At this point the correct answer appears in the window. If the student is correct, he activates the lever in one manner, which eliminates the item from the next sequence. If he is incorrect, the lever is activated in a different manner, thus retaining the item in the next sequence.

Holland (1960), a co-worker of Skinner's, has suggested several well-known learning principles that should be applied to teaching machine technology: immediate reinforcement for correct answers is a must, learned behavior is possible only when it is *emitted* and reinforced, gradual progression (i.e., small steps in learning sequences and reducing wrong answers) is necessary to establish complex repertoires, gradual withdrawal (fading or vanishing) of stimulus support is effective, it is necessary to control the student's observing and echoic behavior and to train for discrimination, the student should write his response. The Skinner machine does in fact employ these principles.

Ferster and Sapon (1958) described the Cardboard Mask, a most simple teaching machine which employs the principles which Skinner and Holland outline so clearly. This device is a cardboard folder containing mimeo-

graphed material which is presented one line at a time. The student, after writing his response on a separate sheet of paper, advances the paper in the mask, thereby exposing the correct response.

In 1958, a number of investigators interested in teaching machines recommended that the programed material be a function of the student's response. This idea suggests that a "wrong" response may not necessarily be negative reinforcement and that both the "right" and "wrong" responses should modify the program. Rath and Anderson (1958) and Rath, Anderson, and Brainerd (1959) have suggested the use of a digital computer which automatically adjusts problem difficulty as a function of the response. Crowder's (1958, 1959a, 1959b) concept of "intrinsic programming" permits the response to alter the programming sequence.

During the last few years, researchers have been focusing their attention on investigating many of the variables which are pertinent to the design and use of teaching machines. The seemingly simple task of defining a teaching machine has been a serious problem to many authors (Day, 1959; Silberman, 1959; Weimer, 1958). Some definitions have made more extensive demands on teaching devices than others. Learning theorists (Kendler, 1959; Porter, 1958; Skinner, 1957; Spence, 1959; Zeaman, 1959) are now most outspoken concerning the application of theoretical concepts to teaching machine technology. Transfer of training, mediational processes, reinforcement, motivation, conditioning, symbolic processes, and language structure are but a few of these areas of interest.

There are indeed many other variables about which there is a diver-

gence of opinion and about which experimental evidence is completely lacking or controversial. The reports of Skinner (1958), Israel (1958), Coulson and Silberman (1960), Fry (1959), and Stephens (1953) are all focused, at least in part, on questions related to response modes, e.g., multiple choice, construction of the response, responses with reinforcement, etc. Briggs, Plashinski, and Jones (1955) investigated self-paced vs. automatically paced machines. The importance of motivation in connection with teaching machines has been explored by Holland (unpublished), Mayer and Westfield (1958), and Mager (1959).

Essentially, the history of automated teaching is short—it started in the mid-twenties and was strenuously reactivated by the appearance of Skinner's 1958 article. Empirical investigations of many important issues in this field are just now beginning to appear. However, the necessity of developing automated teaching methods has been evident for many years.

GENERAL PROBLEM AREAS

Definition

As in any new field, the first problem is one of definition. What is a teaching machine? Silberman (1959) says that a teaching device consists of four units: an input unit, an output unit, a storage unit, and a control unit. As such, this definition includes a broad category of devices, from the most simple to the most complex. Weimer (1958) goes beyond the device itself, stating that a teaching machine must present information to the student as well as test the student by means of a controlled feedback loop. Crowder (1960) insists that a teaching machine

must in some way incorporate two-way communication. That is, the student must respond to the information presented by the machine, and the machine must in turn recognize the nature of the student's response and behave appropriately (p. 12).

Perhaps the most inclusive definition is one given by Day (1959):

A teaching machine is a mechanical device designed to present a particular body of information to the student. . . . Teaching machines differ from all other teaching devices and aids in that they require the active participation of the learner at every step (p. 591).

Although the emphasis in some of the above concepts is different, together they give a rather complete description and, if you will, definition.

Programing

The programing of subject matter for teaching machines is the most extensive and difficult problem in this new technology. Beck (1959) describes specific concepts which he thinks appropriate for programing a Skinner-type machine:

A student's responses may be restricted and guided; in a great number of ways. These range from all types of hints . . . to simply presenting the response which it is desired a student acquire (p. 55).

Carr (1959) discusses in some detail the importance of programing in terms of learning efficiency and retention. Much of what he says remains open for empirical verification. Rothkopf (1960) has suggested that the development of programed instruction suffers from two difficulties: a weak rational basis for program writing and inadequate subject-matter knowledge among program writers.

The extent to which any initial program needs revision is perhaps exemplified by the program in Harvard's course Natural Sciences 114. Holland points out that the first program of materials included 48 disks, each containing 29 frames,

whereas a revision and extension of the program the following year included 60 disks of 29 frames each. Holland's objective was to extend the program and decrease the number of student errors. Crowder's (1960) programing objectives are different from Holland's. He states:

By means of "intrinsic programming" it [the program] recognizes student errors as they occur and corrects them before they can impede understanding of subsequent material or adversely affect motivation (p. 12).

Crowder considers it almost impossible to write a program which completely avoids error, and therefore he structures the program requirements on the probability of error. When an error is made, the next presentation explains the subject's mistake. Depending on the nature of the error and when it occurs, the subject may either return to the original question or enter a program of correctional material.

Another concept for programing is known as *branching* (Bryan & Rigney, 1959). Through branching, many possible routes are provided through which the subject can proceed, depending on the response. The subjects are allowed to skip certain material if they have demonstrated a knowledge of it. One study (Coulson & Silberman, 1960) suggests that under branching conditions subjects require less training time than under nonbranching conditions; however, results on the criterion test were not significantly different.

For certain kinds of subject matter, *vanishing* is still another concept for programing (Skinner, 1958). A complete or nearly complete stimulus is presented to the subject. Subsequent frames gradually omit part of the stimulus until all of it is removed. The subject is then required to reconstruct the stimulus.

To program verbal learning sequences, Homme and Glaser (1959) suggest the Ruleg. With this method, the written program states a rule and provides examples for this rule. In each case, either the rule or the example is incomplete, requiring the subject to complete it.

In a recent study Silverman (1960b) investigated methods of presenting verbal material for use in teaching machines. He recommended that further research involving the design and use of teaching machines should take into consideration the possible use of context cues as a means of facilitating serial rote learning. At the same time, however, he stated that continuous use of context cues as ancillary prompts should be avoided, since such prompts can interfere with learning.

The optimum size of steps and the organization of the programmed material are two formidable problems. Skinner (1958) states:

Each step must be so small that it can always be taken, yet in taking it the student moves somewhat closer to fully competent behavior (p. 2).

In order to determine the value of steps in a program, Gavurin and Donahue (1961) investigated the effects of the organization of the programmed material on retention and rate of learning. They state that the assumption that optimum teaching machine programs are those in which items are presented in a logical sequence has been validated for acquisition but not retention. The results of a study carried out by Coulson and Silberman (1959) indicated that small steps were more time consuming but resulted in statistically significant higher test scores on one of the criterion tests. Pressey (1959) in principle disagrees with Skinner's notions of short and easy steps, and he

strongly suggests an experimental investigation of this question. Both rate of learning and retention (recall or recognition) are of critical concern.

The above discussion suggests several areas which are directly applicable to programing and which are under investigation and/or need further experimentation. Indeed, there are a number of unanswered questions in the programing complex, some of which have been suggested by Galanter (1959):

1. What is the correct order of presentation of material?
2. Is there an optimum number of errors that should be made?
3. How far apart (in some sense) should adjacent items be spaced?
4. Is experimentally controlled pacing more effective (in some sense) than self-pacing?
5. Is one program equally effective for all students?
6. What are the effects of using different programing techniques (branching, intrinsic programing, vanishing) in various subject-matter areas?
7. What criteria are most appropriate in the evaluation of student learning?

These questions are but a few of the intriguing and complex problems facing investigators in the new field of programing material for teaching machines. Answers to these questions will help not only the educator but also the engineer who is concerned with writing adequate specifications for the construction of teaching machines.

Response Mode

The kind of response that should be given by a subject has been a controversial question in the teaching machine field. Pressey's original machine (1926) required the subject to

press a lever corresponding to his choice of answer. The format of the answers was multiple-choice. Skinner (1958) emphasized the necessity of having the subject *compose* (construct) the response. Skinner states:

One reason for this is that we want him to recall rather than recognize—to make a response as well as see that it is right. Another reason is that effective multiple-choice material must contain plausible wrong responses, which are out of place in the delicate process of "shaping" behavior because they strengthen unwanted forms (p. 2).

Coulson and Silberman (1960) investigated this question of multiple-choice vs. constructed response by using *simulated* teaching machines—human beings were used instead of automatic control mechanisms. Their results indicated that the multiple-choice response mode required significantly less time than the constructed response mode and that no significant difference was obtained between response modes on the criterion test. Further, they reported that no significant differences were obtained among the experimental groups on the multiple-choice criterion subtest or on the total (multiple-choice plus constructed response) criterion test. Fry (1959) has discussed this response-mode question along with other variables, and he has carried out extensive research concerning constructed vs. multiple-choice response modes. The results of his study favor the use of constructed response when recall is the objective of the learning.

In addition to the basic controversy (which needs much more investigation) between multiple-choice and constructed responses, there are several "variations on the theme" which are evident. Stephens (1953) has recommended that every wrong answer in a multiple-choice question

appear as a correct choice for another item. He calls this program "inside alternatives." His data indicate that there was no difference between control and experimental groups on a criterion test using either nonsense syllables or Russian unless each right choice appeared as a wrong alternative for the three subsequent items. The use of prompts in general has been shown to be an effective technique in automated teaching (Cook, 1958; Cook & Kendler, 1956; Cook & Spitzer, 1960).

Using learning booklets, Goldbeck (1960) investigated the effect of response mode and learning material difficulty on automated instruction. The three response modes used were: overt response (the subject was required to construct a written response), covert response (the subject was permitted to think of a response), and implicit response (the subject read the response which was underlined). Goldbeck states:

Learning efficiency scores, obtained by dividing quiz scores by learning time, showed that the implicit (reading) response condition produced significantly more efficient learning than the overt response condition. The covert response condition fell between the other conditions in learning efficiency (p. 25).

Concerning quiz-score results, the overt response group

performed significantly poorer than the other response mode groups at the easy level of difficulty. Performance of the overt response group improved significantly at the intermediate difficulty level to the extent that it exceeded the performance of all other groups (pp. 25–26).

Goldbeck concludes that

doubt is cast upon the assumption that the best learning is achieved by use of easy items and requiring written constructed responses (p. 26).

To the author's knowledge, the use of an oral response in conjunction

with the Skinner teaching machine and its effect on learning rate and retention have not been reported in the literature. Furthermore, the importance of response mode as a function of reinforcement must be specified. Israel (1958) has suggested that natural and artificial reinforcement may affect the subjects' learning. A most comprehensive analysis of response-mode and feedback factors has been reported by Goldbeck and Briggs (1960).

The general area of reinforcement suggests problems related to the drop-out feature of teaching machines. Pressey's (1927) original machine dropped items after the correct answer had been given twice. Skinner's machines at the Harvard Psychological Laboratory also have the drop-out feature, although the commercially available machines based on Skinner's design do not incorporate this feature. With reference to a study carried out at Harvard, Holland (unpublished) reported significantly superior performance when the drop-out feature was used.

If items are dropped, the sequence of items is of course changed. How important is the sequence? If items should be dropped, by what criterion of learning can one justify omitting an item from the sequence? If items are not dropped and the criterion for the learning procedure is a complete run (i.e., once through the sequence without error), what is the effect upon retention? Being correct is positive reinforcement; thus, some items under these circumstances will receive a greater amount of positive reinforcement than others. What would be the effect of additional reinforcements with or without drop-out? Again, a plethora of problems and a paucity of answers!

Response time, another important

variable, has been investigated by Briggs, Plashinski, and Jones (1955). Their study suggests that there is no difference between self-paced and automatically paced programs as determiners of response time. However, the problem of pacing for individual items is still a recent one and needs further research. Another aspect of response time—the distribution of practice—has been studied extensively since Ebbinghaus' investigation in 1885. For example, Holland (unpublished) states that in an experiment at Harvard "a few students completed all the disks in a small number of long sessions while others worked in many short sessions. . . . Apparently the way practice was distributed made little difference" (p. 4). Nevertheless, the distribution of practice, like the problem of pacing, is yet a subject of controversy, with most investigations favoring some form of distributed practice (Hovland, 1951).

The above section outlines briefly some of the major problems associated with the variables affecting response mode. Although some of the variables have already been investigated, these and others, together with their interactions, need further research.

Knowledge of Results

There are many peripheral problems related to teaching machines, one of which is the effect of immediate knowledge of results on learning. Angell (1949), using a multiple-choice punchboard technique, found that "learning is significantly enhanced by immediate knowledge of results." Briggs (1949), also using the Punchboard, confirmed these results. Bryan and Rigney (1956) noted superior performance when subjects were given knowledge of results, specifically, an explanation

if the answer was incorrect. This last study was later expanded by Bryan, Rigney, and Van Horn (1957), who investigated differences between three kinds of explanation given for incorrect response. None of the three types of explanation proved to be superior in teaching the subjects. Because of their controvertible results, the above studies demonstrate that, although immediate knowledge of results appears to be effective in the learning process, this problem contains many facets which need more empirical data.

Motivation

One of the many reasons given for the effectiveness of teaching machines is that the student's motivation is increased. Psychologists and educators have realized for some time that the motivation variable ranks very high among those variables pertinent to learning. In 1958 and 1959, Holland surveyed the use of the teaching machine in classes at Harvard. He found that most students felt that they would have gotten less out of the course if the machines had not been used, that most students preferred to have machines used for part of the course, and finally that most students felt that the teaching machine was used by the instructor "to teach me as much as possible with a given expenditure of my time and effort." During a field tryout of the Subject-Matter Trainer in the Semi-automatic Ground Environment System, Mayer and Westfield (1958) observed that "motivation to work with the trainer is high." The supervisory as well as the operational personnel encouraged the use of this training technique.

Mager (1959) suggests that motivation and interest are a function of the percentage of correct responses.

He observed that in two young subjects negative feelings for learning mathematics in the usual classroom situation did not transfer to learning mathematics by means of a teaching machine. The cause of this phenomenon is perhaps best explained by the subjects' statement that, because they were able to understand the programed material, it did not seem to be mathematics at all. This interesting relationship between comprehension and motivation needs further investigation.

Equipment

There are many inexpensive models of teaching machines which will soon hit the consumer market. For much of this equipment, there is very little experimental evidence which supports the various designs. As previously pointed out, Holland has collected data which support the efficiency of the drop-out feature in a teaching machine; yet commercial models presently available do not incorporate this feature, presumably because of its high cost. Generally, it seems that production is now and will continue to be out of phase with much of the research which has provided necessary teaching machine specifications. Moreover, because of their expense, it is likely that some very important features will be omitted in manufacture.

The methods of displaying programed material, another unexplored problem area, must be investigated so as to provide the design engineer with requirements based on empirical findings. The display problem is less acute, perhaps, with material for the elementary school than it is with programs designed to teach maintenance procedures and aspects of the biological sciences.

The use of computer controlled

teaching machines has been recommended by many authors (Coulson & Silberman, 1959; Skinner, 1958). Utilizing a central computer, with many programs capable of adapting to individual needs and of providing stimulus materials to 50 or more students simultaneously, is a feasible notion for large-scale training programs. With a computer, the display problem again becomes a major issue. Training in pattern recognition, information handling, and display interpretation are but a few appropriate areas which should be studied. The alternate modes of presentation become more extensive as computer capacity increases. In the case of certain kinds of subject matter, a computer generated, pictorial display of information may be a more effective presentation than other display techniques. Future research must solve these problems in equipment design.

Teaching Machines and Other Techniques

The use of automated teaching devices may be optimized, perhaps, if there is a proper balance between this technique and other compatible teaching methods. What percentage of a course should be machine taught? What subject matter is best suited to automated devices? If classroom courses were as carefully and thoughtfully programed as some of the programs currently being prepared for teaching machines, might some of the advantages of machines diminish? Perhaps some of the apparent advantages of teaching machines are no more than methods of illustrating correctable classroom techniques! It might well be that the instructor's enthusiasm and inspiration, a factor supposedly dominant in higher education, is vital in mastering a particular subject-matter area. Will

creativity in certain students be harmed by extensive education via the machine? Again, consideration of the use of a teaching machine, the subject matter, the program, the level of education, and the techniques used in combination with the teaching machine provide a fertile field for experimentation. As of now, questions in this area remain unanswered. Silverman (1960a) has presented an excellent, detailed discussion of problems inherent in this new technology of automated teaching and the current trends in the field.

PROBLEMS OF APPLICATION

The most obvious problems in the attempt to use automated teaching techniques have been outlined in the previous section. There is still much of the unknown associated with techniques, machines, programing, etc. to be eliminated before a direct solution to a particular training problem can be specified. Many alternatives exist, the best of which has not yet been determined. In addition to these voids, there is a serious lack of definition in the objectives of many training programs.

What is the objective of a particular automated course or program? From a pragmatic point of view, what are the criteria by which a specific educational program can be evaluated? For example, the objectives might range from the teaching of rote tasks to the presentation of more abstract material. Needless to say, the techniques for both teaching and evaluating learning could be substantially different in each case. The purpose of teaching, the objective of an educational program, must be initially defined. Only then will the concepts *learning* and *teaching* be meaningful in a particular context.

After definition, the next step is to determine what subject matter will

provide the student with the necessary information. It is at this point that the major pitfall in education is likely to appear. Even though many training programs do not have a defined objective, their course content is nonetheless prescribed, and the text and/or materials used in previous, nonautomated courses become the prime source of material for an automated teaching program. To program an automated teaching machine with presently available materials might well result only in a more efficient method of teaching the wrong material!

The third step requires decisions in the selection of appropriate teaching techniques. Answers to questions involving programing, choice of teaching machine, learning procedures, pacing, and response modes are still not known.

The fourth and last step requires an evaluation of the selected automated teaching method in terms of

the originally established objectives. Conventional methods of instruction should be compared with the innovative methods by means of a specific set of criteria, e.g., in terms of training time, job performance, retention of learned information, etc.

The questions confronting the researcher in teaching machine technology are one example of the broader questions of man-machine interrelation. Data pertinent to the principles of human engineering, the optimum man-machine interaction, the degree to which the machine can perform functions formerly allocated to man, and the appropriate allocation of functions between man and machine will be provided by a research program investigating teaching machines. Inadequate attention to any of the above-mentioned steps will result in failure to provide the needed answers in a field which may increase training effectiveness and reduce training costs.

REFERENCES

- ANGELL, G. W. The effect of immediate knowledge of quiz results on final examination scores in freshman chemistry. *J. educ. Res.*, 1949, 42, 391-394.
- ANGELL, G. W., & TROYER, M. E. A new self-scoring test device for improving instruction. *Sch. Soc.*, 1948, 67, 84-85.
- BECK, J. On some methods of programming. In E. H. Galanter (Ed.), *Automatic teaching: The state of the art*. New York: Wiley, 1959. Pp. 55-62.
- BESNARD, G. G., BRIGGS, L. J., MURSH, G. A., & WALKER, E. S. Development of the subject-matter trainer. *USAF Personnel Train. Res. Cent. tech. Memo.*, 1955, No. ASPRL-TM-55-7.
- BESNARD, G. G., BRIGGS, L. J., & WALKER, E. S. The improved subject-matter trainer. *USAF Personnel Train. Res. Cent. tech. Memo.*, 1955, No. ASPRL-TM-55-11.
- BRIGGS, L. J. The development and appraisal of special procedures for superior students and an analysis of the effects of knowledge of results. *Abstr. Doctoral Dissertations, Ohio State U.*, 1949, No. 58.
- BRIGGS, L. J., PLASHINSKI, D., & JONES, D. L. Self-pacing versus automatic pacing of practice on the subject-matter trainer. *USAF Personnel Train. Res. Cent. lab. Note*, 1955, No. ASPRL-LN-55-8.
- BRYAN, G. L., & RIGNEY, J. W. An evaluation of a method for ship-board training in operations knowledge. *U. Sth. Calif. Electronics Personnel Res. Group tech. Rep.*, 1956, No. 18.
- BRYAN, G. L., & RIGNEY, J. W. Current trends in automated tutoring and their implications for naval technical training. *U. Sth. Calif. Dept. Psychol. tech. Rep.*, 1959, No. 29.
- BRYAN, G. L., RIGNEY, J. W., & VAN HORN, C. An evaluation of three types of information for supplementing knowledge of results in a training technique. *U. Sth. Calif., Electronics Personnel Res. Group tech. Rep.*, 1957, No. 19.
- CARR, W. J. Self-instructional devices: A review of current concepts. *USAF WADC tech. Rep.*, 1959, No. 59-503.
- COOK, J. O. Supplementary report: Processes underlying learning a single paired-associate item. *J. exp. Psychol.*, 1958, 56, 455.

- COOK, J. O., & KENDLER, T. S. A theoretical model to explain some paired-associate learning data. In G. Finch & F. Cameron (Eds.), *Symposium on Air Force human engineering, personnel, and training research*. Washington, D.C.: National Academy of Sciences-National Research Council, 1956. Pp. 90-98.
- COOK, J. O., & SPITZER, M. E. Supplementary report: Prompting versus confirmation in paired-associate learning. *J. exp. Psychol.*, 1960, 59, 275-276.
- CORRIGAN, R. E. Automated teaching methods. *Automated teach. Bull.*, 1959, 1(2), 23-30.
- COULSON, J. E., & SILBERMAN, H. F. Proposal for extension of automated teaching projects. *Sys. Developm. Corp. field Note*, 1959.
- COULSON, J. E., & SILBERMAN, H. F. Effects of three variables in a teaching machine. *J. educ. Psychol.*, 1960, 51, 135-143.
- CROWDER, N. A. *An automatic tutoring book on number systems*. Vol. 1. Timonium, Md.: Hoover Electronics Co., 1958.
- CROWDER, N. A. Automatic tutoring by means of intrinsic programming. In E. H. Galanter (Ed.), *Automatic teaching: The state of the art*. New York: Wiley, 1959. Pp. 109-116. (a)
- CROWDER, N. A. The concept of automatic tutoring. *USAF Personnel Train. Res. Cent. organizational Pap.*, 1959. (b)
- CROWDER, N. A. The "tutor." *J. Amer. Soc. Train. Dir.*, 1960, 14(5), 12-17.
- DAY, J. H. Teaching machines. *J. chem. Educ.*, 1959, 36, 591-595.
- FERSTER, C. B., & SAFON, S. M. An application of recent developments in psychology to the teaching of German. *Harv. educ. Rev.*, 1958, 28, 58-69.
- FRY, E. B. Teaching machine dichotomy: Skinner versus Pressey. Paper presented at American Psychological Association, Cincinnati, September 1959.
- FRY, E. B., BRYAN, G. L., & RIGNEY, J. W. Teaching machines: An annotated bibliography. *Audio-Visual commun. Rev.*, 1960, 8, Suppl. 1, 1-80.
- GALANTER, E. H. The ideal teacher. In E. H. Galanter (Ed.), *Automatic teaching: The state of the art*. New York: Wiley, 1959. Pp. 1-11.
- GAVURIN, E. I., & DONAHUE, V. M. Logical sequence and random sequence. *Automated teach. Bull.*, 1961, 1(4), 3-9.
- GOLDBECK, R. A. The effect of response mode and learning material difficulty on automated instruction. *Amer. Inst. Res. tech. Rep.*, 1960, No. AIR-328-60-IR-124.
- GOLDBECK, R. A., & BRIGGS, L. J. An analysis of response mode and feedback factors in automated instruction. *Amer. Inst. Res. tech. Rep.*, 1960, No. AIR-328-60-IR-133.
- HOLLAND, J. G. Teaching machines: An application of principles from the laboratory. In *Proceedings of the Educational Testing Service Invitational Conference*, October 1959, *The impact of testing on the educational process*. Princeton, N. J.: Educational Testing Service, 1960.
- HOMME, L. E., & GLASER, R. Problems in programing verbal learning sequences. Paper presented in symposium on research issues in study of human learning raised by developments in automated teaching methods, American Psychological Association, Cincinnati, September 1959.
- HOVLAND, C. I. Human learning and retention. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 613-689.
- ISRAEL, M. L. Skinnerian psychology and educational redesign. Paper read in symposium, American Psychological Association, Washington, D. C., September 1958.
- KENDLER, H. H. Teaching machines and psychological theory. In E. H. Galanter (Ed.), *Automatic teaching: The state of the art*. New York: Wiley, 1959. Pp. 177-185.
- LITTLE, J. K. Results of use of machines for testing and for drill upon learning in educational psychology. *J. exp. Educ.*, 1934, 3, 45-49.
- MAGER, R. F. Preliminary studies in automated teaching. *IRE Trans. Educ.*, 1959, E-2, 104-107.
- MAYER, S. R., & WESTFIELD, R. L. A field tryout of a teaching machine for training in SAGE operations. *USAF Cambridge Res. Cent. tech. Memo.*, 1958, No. OAL-TM-58-16.
- PETERSON, J. C. A new device for teaching, testing, and research in learning. *Trans. Kans. Acad. Sci.*, 1930, 33, 41-47.
- PORTER, D. Teaching machines. *Harv. Grad. Sch. Educ. Ass. Bull.*, 1958, 3(1), 1-5.
- PRESSEY, S. L. A simple apparatus which gives tests and scores—and teaches. *Sch. Soc.*, 1926, 23, 373-376.
- PRESSEY, S. L. A machine for automatic teaching of drill material. *Sch. Soc.*, 1927, 25, 549-552.
- PRESSEY, S. L. A third and fourth contribution toward the coming "industrial revolution" in education. *Sch. Soc.*, 1932, 36, 668-672.
- PRESSEY, S. L. Development and appraisal of devices providing immediate automatic scoring of objective tests and concomitant

- self-instruction. *J. Psychol.*, 1950, 29, 417-447.
- PRESSEY, S. L. Certain major psycho-educational issues appearing in the conference on teaching machines. In E. H. Galanter (Ed.), *Automatic teaching: The state of the art*. New York: Wiley, 1959. Pp. 187-198.
- RAMO, S. A new technique of education. *Engng. sci. Mon.*, 1957, 21 (October), 17-22.
- RATH, G. J., & ANDERSON, NANCY S. The IBM research center teaching machine project: I. The teaching of binary arithmetic. II. The simulation of a binary arithmetic teaching machine on the IBM 650. Paper presented at USAF Office of Scientific Research symposium on teaching machines, University of Pennsylvania, December 8-9, 1958.
- RATH, G. J., ANDERSON, NANCY S., & BRAINERD, R. C. The IBM research center teaching machine project. In E. H. Galanter (Ed.), *Automatic teaching: The state of the art*. New York: Wiley, 1959. Pp. 117-130.
- ROTHKOPF, E. Z. A do-it-yourself kit for programmed instruction. *Teach. Coll. Rec.*, 1960, 62, 195-201.
- SILBERMAN, H. F. Introductory description of teaching machines (physical characteristics). Paper read in symposium on automated teaching, Western Psychological Association, San Diego, California, April 1959. (Abstract)
- SILVERMAN, R. E. Automated teaching: A review of theory and research. USN Training and Development Center, 1960. (ASTIA AD-241 283) (a)
- SILVERMAN, R. E. The use of context cues in teaching machines. USN Training and Development Center, 1960. (ASTIA AD-238 777) (b)
- SKINNER, B. F. The science of learning and the art of teaching. *Harv. educ. Rev.*, 1954, 24, 86-97.
- SKINNER, B. F. *Verbal behavior*. New York: Appleton-Century-Crofts, 1957.
- SKINNER, B. F. Teaching machines. *Science*, 1958, 128, 969-977.
- SPENCE, K. W. The relation of learning theory to the technology of education. *Harv. educ. Rev.*, 1959, 29, 84-95.
- STEPHENS, A. L. Certain special factors involved in the law of effect. *Abstr. Doctoral Dissertations, Ohio State U.*, 1953, No. 64.
- WEIMER, P. K. A proposed "automatic" teaching device. *IRE Trans. Educ.*, 1958, E-1, 51-53.
- ZEAMAN, D. Skinner's theory of teaching machines. In E. H. Galanter (Ed.), *Automatic teaching: The state of the art*. New York: Wiley, 1959. Pp. 167-176.

(Early publication received March 5, 1961)

DEVELOPMENT OF RESEARCH ON THE PHYSIOLOGICAL MECHANISMS OF AUDITORY LOCALIZATION

MARK R. ROSENZWEIG
University of California, Berkeley

How does a listener determine the direction from which a sound comes? In the last century, physicists and physiologists usually gave a psychological explanation; they held that the listener makes a judgment by comparing the differing stimulation at the two ears. In the present century psychologists and physiologists have been seeking a physiological explanation; they are attempting to find where and how nerve impulses originating at the two ears interact in the brain. In this paper we will trace the development of hypotheses concerning the mechanisms of auditory localization. ("Localization" will signify here the perception of the *direction* of a sound source; perception of *distance* will not be considered.)

First, however, let us note briefly the importance of auditory localization in animal and human life. Some animals owe their livelihood to their ability to localize. Thus, certain types of bat catch their insect prey in the dark by echolocation, and certain moths, in turn, attempt to detect and avoid these bats by auditory clues (Griffin, 1958). Human beings also have considerable ability to localize sounds. Localization provides the basis for the detection of obstacles by blind people, an ability which long remained a mystery under the ambiguous designation of "facial vision of the blind." In the 1940s Cornell psychologists proved conclusively that this performance depends on localization of echoes reflected from the obstacles (Cotzin & Dallenbach, 1950; Supa, Cotzin, & Dallenbach, 1944).

Seeing persons, too, benefit greatly from their ability to localize sounds. Localization makes it easier to listen to one signal or message in the presence of competing signals—a task which communication engineers refer to as "the cocktail party problem" (Cherry, 1957). Because sounds can be discriminated more readily when they are heard as coming from different directions, binaural hearing aids and stereophonic recordings of music are becoming steadily more popular.

INITIAL WORK ON THE MECHANISMS OF AUDITORY LOCALIZATION

The first person to investigate the nervous system in regard to localization seems to have been Louis Jurine, a naturalist, anatomist, and physician of Geneva. Lazaro Spallanzani of Pavia had demonstrated in 1793 that blinded bats could fly and avoid obstacles just as well as normal bats, but he could not imagine what sense then substituted for vision (1798). Jurine took up the question and decided that the solution must lie "at the tip of a scalpel." Noting the large size of the external ear of the bat, he went on to find that "a considerable neural apparatus" was devoted to hearing. Unfortunately, the published extract of his account (1798) does not give any fuller indication of Jurine's neuroanatomical findings. (The central connections of the auditory nerve seem not to have been discovered until almost a century later. In the bat the cochlear nucleus bulges out from the medulla right to the tip of the cochlea, and it is possi-

ble that Jurine took this for a large auditory nerve.) Following this lead, Jurine devised ingenious behavioral tests which provided clear evidence that the blinded bat guides itself by auditory clues.

Venturi and the Intensity Hypothesis

The first experiments on auditory localization by human observers seem to be those reported in two similar articles (1800a, 1800b) by Giovanni Battista Venturi, Professor of Physics at Modena and Pavia.¹ Venturi's experiments were similar to those that Lord Rayleigh performed independently 75 years later (1877). Venturi concluded that "The inequality of the two simultaneous sensations of the two ears informs us of the true direction of the sounds" (1800a, p. 386). Part of his evidence was that sounds coming from directly in front of the observer could not be distinguished from sounds coming directly from the rear, if the observer kept his head still. Venturi also con-

cluded that a person with one deaf ear must turn his head to localize and will usually make errors in localizing sounds that are very brief. (As we shall see, certain recent experimenters have ignored the role of head movements in localization.)

Venturi noted that philosophers had attempted to explain the singleness of vision by the convergence of the two optic nerves, but he held that this was not the case for hearing:

Since we distinguish the two simultaneous sensations of the two ears, and since their different intensities furnish us knowledge of the true direction of the sound, therefore one must conclude that the two sound impressions do not mix together inside the skull (1800a, p. 388).

Venturi furthermore concluded that the visual impressions of the two eyes do not mix, citing the phenomenon that was later to be called "retinal rivalry."

Venturi's intensity theory remained the dominant explanation for localization until early in the twentieth century. It was propounded, for example, by Magendie (1831) and Johannes Müller (1840). Magendie offered it as something evident and did not credit its discovery to anyone. Müller claimed that perception of direction of sound "is an act of judgment which founds it on experience previously acquired. . . . The only true guide for this inference is the more intense action of the sound upon one than upon the other ear" (p. 479). He further noted that when a sound comes from directly ahead or behind, it falls equally upon the two ears and is then impossible to localize; this demonstration he ascribed to "Venturini" (sic).

PROGRESS IN THE SECOND HALF OF THE NINETEENTH CENTURY

The first person to abandon the judgmental interpretation of localiza-

¹ This work of Venturi seems to have set the style for most of the research on auditory localization during the nineteenth century. Yet, while some of his findings soon became common knowledge, it was forgotten who had discovered them. Thus Klemm, in his detailed history of auditory localization (1914), mentioned Venturi in a single sentence and only in regard to effects of head movements. Pierce (1901) and Boring (1942) did not mention Venturi in their historical discussions of localization; both ascribed the first experiment on localization to E. H. Weber in 1848.

Venturi's first publication on this subject is even earlier than those indicated above. A paper of 1796 in French gives almost the same material as that of the German articles of 1800 (see Venturi, 1796).

In 1801 Venturi published a report in Italian on this research, appending it to the second edition of his book on physical research on color (see Venturi, 1801). The contents of this version are similar to those of the German articles. Thus it appears that Venturi made repeated attempts to secure a wide public for this research.

tion seems to have been S. Scott Alison, an English physician. He had invented the "differential stethophone" which consisted simply of two stethoscopes, one for each ear. After using this instrument, Alison reported that a sound is restricted to the ear that receives it in greater intensity and is suppressed in the other ear. This, he remarked,

holds apparently in virtue of a law seemingly established for the purpose of enabling man and the lower animals to determine the direction of the same sound, with more accuracy than could be done had a judgment to be formed between the intensity of two similar sensations in the two ears respectively (1858, pp. 388-389).

Alison was considerably ahead of his time in this formulation but his work seems to have had little influence.

Sylvanus P. Thompson wondered about the basis of binaural beats, heard when he connected two slightly mistuned forks, one to each ear (1877). (Lord Rayleigh was to observe later—1907—as Dove—1857—had earlier, that the sound also changed location while beating.) Thompson rejected the hypothesis that bone conduction could account for binaural beats. Noting that the auditory nerves do not decussate as the optic nerves do, he concluded "that any means of comparison which may exist in the nerve systems of the ears exists deep-seated in the actual structure of the brain" (1877, p. 276). In his next paper he again noted that it is problematical where the sensations from the two ears "blend," and he remarked, "This point deserves the attention of anatomists and physiologists" (1878, p. 389).

Tracing the Afferent Auditory Pathways

Physiologists and anatomists had, in fact, already turned their attention to the localization of sensory proc-

esses in the brain. The study of cerebral localization of function had recently been given new impetus by the introduction of a new method—precise electrical stimulation, introduced by Fritsch and Hitzig in 1870. The revolutionary results obtained by this method led also to renewed interest in experiments involving precise ablation, as a check on the electrical experiments. David Ferrier, Professor of Neuropathology in London, began extensive mapping of the brain of several species in 1873, using electrical and surgical techniques (1890). In the superior temporal convolution of monkeys and its homologues in other species, electrical stimulation produced the same reaction as if a shrill sound had been made in the contralateral ear. The animal pricked up or retracted the ear and often moved its head or eyes to that side. (Much later the same experiment was to be performed in conscious surgical patients—Penfield & Rasmussen, 1950. In most cases, the human subjects reported hearing sounds on the side contralateral to the stimulated hemisphere; some sounds were heard "bilaterally"; no sounds were heard ipsilaterally.) Ferrier also claimed that ablation of the auditory cortex of both hemispheres made monkeys inattentive to sound. Heschl in 1878 succeeded in tracing the auditory tracts to the superior temporal convolution, and the auditory area is often given his name.

Luigi Luciani of Florence used the ablation technique extensively in cortical mapping (Luciani, 1884; Luciani & Seppilli, 1886; Luciani & Tamburini, 1879). One behavioral test devised by Luciani has recently been reintroduced by Riss (1959). In this test, bits of food were thrown to the floor near the blindfolded animal, and the accuracy of its reactions to the sounds was noted. Luciani con-

firmed that the auditory area of the cortex is located posteriorly in the temporal lobe. He found that

each ear [has] connections with both auditory spheres, but chiefly with that of the opposite side. In fact, every unilateral extirpation of sufficient extent in the province of the auditory sphere causes a bilateral disorder of hearing, more marked on the opposite side . . . (1884, p. 155).

Thus, after an ablation in the right hemisphere, these results were reported by Luciani and Seppilli:

Hearing is affected on both sides, but more at the left than at the right ear. The animal shows that it hears the sound of pieces of food falling to the left, but it mistakes the direction and turns to the other side. At the right ear this does not happen (1886, p. 79).

The effects of unilateral lesions generally disappeared within a few weeks; they persisted longer the larger the lesion. None of the lesions covered the whole auditory area as it is now defined, so it cannot be told from these studies whether a complete unilateral lesion would have led to some permanent impairment of localization. Bilateral lesions of the auditory areas were found to produce permanent perceptual impairment, as this example indicates:

When called suddenly, the dog reveals through its movement that it is not deaf; but it does not follow and does not turn its head toward the sound, but, in fact, often even turns to the other side; in short, it seems not to understand what it hears and not to perceive the direction of sounds (psychic deafness) (1886, p. 119).

Luciani concluded that in the auditory system, just as in the visual system

we must distinguish a crossed and a direct fasciculus; the former consisted of a much larger number of fibers than the latter. Neither of these fasciculi possesses any uniform relation with distinct segments of their respective cortical spheres, but their fibers irradiate themselves throughout the area of these centres (1884, p. 155).

Histological degeneration studies of

Baginsky, Flechsig, and von Monakow soon revealed rather completely the course of the auditory pathway, crossing in part in the brainstem and proceeding by way of the inferior colliculus and medial geniculate body to the cortex of the temporal lobe (Ferrier, 1890).

After these early achievements, progress not only lagged but some of the findings were even forgotten by workers in the field. Thus, for example, it was taken as surprising in 1928 when removal of one cerebral hemisphere of a patient did not destroy hearing in the opposite ear (Bunch, 1928).

THE EARLY TWENTIETH CENTURY: THE PHASE HYPOTHESIS

Although the role of dichotic phase differences in auditory localization had been shown by Dove (1857), Thompson (1877), and others in the nineteenth century, the phase hypothesis was firmly established only in the twentieth century. Lord Rayleigh had considered this hypothesis previously, but his advocacy of it in 1907 convinced others. Rayleigh showed that while dichotic intensity differences permit localization of high frequency sounds, it is dichotic phase differences that permit localization of low frequency sounds. This belated recognition of the dual basis of localization might have been expected to embarrass supporters of the judgmental position. In fact, Rayleigh did remark, "Perhaps it is not to be expected that we should recognize intuitively the very different basis upon which our judgment rests in the two cases" (1907, p. 203). Nevertheless he never abandoned the judgmental interpretation of localization.

Bowlker (1908), who experimented on the role of phase differences,

speculated briefly about neural interaction:

we may suppose that the transmission of sound impulses through some specialized part of the auditory apparatus or brain takes a definite time from each ear, and that the point where the impulses meet is the focus that gives rise to the sensation of a sound image (p. 327).

From the vantage point of the present, this seems to anticipate Jeffress' hypothesis of 1948, but Bowlker's suggestion is so terse that we cannot be sure.

The success of the Stenger test for unilateral auditory malingering (1907) could have been taken as evidence against the judgmental approach to localization, but it does not seem to have been. The test, recently termed "the most reliable and effective of all malingering tests" (Watson & Tolan, 1949), works in the following way: A subject who simulates deafness of one ear will report hearing a tone that is delivered only to his other ear. When the tone is next delivered to both ears, and more intensely to the ear whose deafness is feigned, the subject will hear it only at this supposedly deaf ear. The malingerer will therefore give himself away by reporting that he does not hear the sound, in spite of the fact that it is present in audible intensity at the admittedly good ear. Thus, it was clear to Stenger (and it should have been to all users of his test) that the listener hears only a single sound and does not compare separate sensations arising from the two ears. This advance in clinical testing had no apparent influence on the development of thinking about auditory localization.

THE TIME HYPOTHESIS

The hypothesis that dichotic time provides a basis for localization seems

first to have been proposed seriously by Mallock (1908) and first to have been demonstrated experimentally by Aggazzotti (1911). It was brought to wide attention in several publications at the end of the first World War (Klemm, 1918, 1920; Piéron, 1922; von Hornbostel & Wertheimer, 1920). Dichotic stimuli separated by as little as 30 microseconds were shown to be perceived toward the side of the prior component. The time hypothesis is incompatible with the judgmental approach, for the dichotic time intervals which provide for localization lie under the threshold of fusion of successive auditory stimuli. That is, dichotic stimuli with a time interval less than 2 milliseconds give rise to perception of a single localized auditory event; there are not two perceptual events that can be compared in order to judge localization. Perhaps the first to recognize the incompatibility of the time hypothesis and the judgmental approach were Kreidl and Gatscher (1923). Their conclusion was to reject the time hypothesis! Since they showed that stimuli must be separated by about 20 milliseconds to be judged as successive, they denied that smaller intervals could have any effect in perception. von Hornbostel (1926) showed the fallacy of this argument. Furthermore, any observer who attempted to test the time hypothesis could verify it. The success of the time hypothesis thus helped to overcome the judgmental approach and to clear the way for work on the physiological mechanisms of localization.

Hypothesized Central Mechanisms

After the role of time differences was demonstrated, several further hypotheses about the mechanisms of localization were soon proposed. von

Hornbostel (1926) suggested that intensity differences were converted into time differences in the auditory system, a stronger stimulus evoking neural responses with less latency than a weaker stimulus. Kemp and Robinson (1937b) were able to demonstrate that the latency of auditory impulses does decrease with intensity, but only within 40 db. of threshold. Stevens and Davis in their book, *Hearing, Its Psychology and Physiology* (1938), mentioned the work of Kemp and Robinson and concluded that the effect of intensity differences cannot result solely from changes in latency. Their reason was that changes in the binaural intensity ratio can shift the location of intense tones. In fact, the change of ratio had been found to be smallest when the tone was about 80 db. above threshold (Upton, 1936); at this level there is no longer a change in latency, according to Kemp and Robinson's results. (This point, we may note, is all that Stevens and Davis had to say about the physiological correlates of auditory localization.) Later research (e.g., Pestalozza & Davis, 1956) has shown that latency continues to decrease with intensity up to at least 70 db. above threshold; this gives new support to von Hornbostel's hypothesis.

Boring (1926) suggested that the locus of cortical excitation might be the physiological correlate of localization. He hypothesized that the ears project, in each cerebral hemisphere, to cortical areas that are not coincident but which overlap. If one ear was stimulated either earlier or more strongly than the other, then the cortical excitation would be located mainly in the projection area of that ear.

Trimble (1928) proposed a vague central hypothesis in which the inter-

aural differences are transmitted to the cortex where localization occurs.

The directional localization of a sound source, under ordinary conditions of hearing, depends upon the configurational nature of the cortical effects that correspond to the physical "difference-pattern" at the ears (p. 523).

von Békésy (1930) proposed a rather detailed schema. He pictured a region of cells where the auditory tracts from the two ears join. Auditory localization would depend upon the proportions of the region that each side excited. Both greater intensity and prior arrival would favor the ear so stimulated.

Woodworth described a possible mechanism in the discussion of auditory localization in his text, *Experimental Psychology* (1938):

It must be a unitary mechanism capable of turning the head in either direction and responsive to nerve currents from both ears. When the currents arrive from both ears, but more from one ear, that side has the advantage. When the current from one ear arrives at the central mechanism ahead of the other and gets in its work first, it has the advantage (p. 533).

Jeffress (1948) suggested a neural "mechanism for the representation of a time difference as place" in the auditory system. He pictured a center where tracts from both ears make common synaptic connections. Within this center there are places where the conduction time is slightly longer from one ear than from the other. If the two ears are stimulated simultaneously, the impulses meet and summate at the locus where the conduction times from both sides are equal. If one ear is stimulated before the other, then the impulses meet at a different locus—a locus where the difference in conduction times compensates for the dichotic time difference. Intensity difference is translated into time difference and this is

then handled in the same way. Jeffress ventured that this mechanism might be located in the medial geniculate, relying on the electrophysiological evidence of Kemp and Robinson (1937a) that no binaural interaction could be found at the lateral lemniscus. Later Jeffress disavowed this location and suggested that if his hypothesized mechanism exists, it exists in the accessory nucleus of the superior olive (1958).

LOCALIZATION INVESTIGATED BY PHYSIOLOGICAL TECHNIQUES

While these hypothetical mechanisms were being proposed, further physiological findings were being obtained by both ablation and electrophysiological techniques.

Information from Ablation Studies

Pavlov (1927) reported a finding of Bikov that a dog could not learn to discriminate between right and left positions of a sound source after the corpus callosum had been transected (p. 150). Girden (1940) found, on the contrary, that dogs retained a learned right-left discrimination of a tone or bell after transection of the corpus callosum. Girden's animals were trained to respond by flexing a leg when the sound came from one side and not to respond when it came from the other. Each sound lasted for 2 seconds. They also retained the discrimination after hemidecortication, losing it only after complete bilateral ablation of the auditory cortex. Even after bilateral ablation, the dog could still orient itself to sound, but it could not be retrained to show the conditional discrimination.

ten Cate (1934) reported that decorticate cats could orient to a sound stimulus. He used various sounds, all of them lasting for 15 to 20 seconds. The physiologists Bard and

Rioch (1937) also reported that decorticate cats could localize sounds accurately, but they did not offer any quantitative observations.

Measurement of impaired ability to localize by cats with bilateral ablations of the auditory cortex was furnished by Neff and his collaborators (see Neff & Diamond, 1958, for a history of several stages of this research). Cats were trained to go to the one of two boxes behind which a buzzer sounded. Intact animals could do this when the boxes were only 5 degrees apart, the angle being measured from the point at which the cats were released into the test area. Animals with complete bilateral destruction of the auditory cortices could discriminate only when the boxes were 40 degrees apart. Three different hypotheses were advanced, any of which might account for the observed results: (a) "An intact auditory cortex is necessary in order that the relationship between auditory signal and food reward may be learned" (Neff, Fisher, Diamond, & Yela, 1956, p. 510). Further experimentation refuted this hypothesis, since cats with bilateral ablation of auditory areas could learn to open a single door when a buzzer sounded. (b) "An intact auditory cortex is essential for maintenance of attention to an auditory signal." (c) "An intact auditory cortex is necessary for accurate localization of sound in space" (1956, p. 511). Neff and Diamond (1958) also reported preliminary results indicating that ability to localize in their tests

is not affected by section of the corpus callosum, is affected very little if at all by section of the commissure of the inferior colliculus, but is severely affected by section of the trapezoid body (p. 108).

Riss (1959) noted that auditory signals of relatively long duration had

been employed in the studies of ten Cate and Neff. He raised the question whether *binaural* localization was actually tested in their work, since it is well known that monaural localization is possible if head movements can be performed while the sound continues. (It will be remembered that Venturi had made this point in 1800.) Riss pointed out that animals of Bard and Rioch oriented to the stimulus slowly, using noticeable head and ear movements. Riss therefore employed both very brief sounds and sounds lasting as long as 30 seconds in experiments with cats. For brief sounds, bits of food were thrown down beside the cat; this is the method that Luciani had used 75 years previously, although Riss evidently did not know of his work. The results replicated and extended those of Luciani. Cats with control lesions were successful with both types of signal. Cats with bilateral ablation of the auditory areas "showed evidence of being unable to orient to brief sounds but were partially successful in seeking out the region of the sound if the sound was prolonged" (p. 383). Tests revealed that these animals could maintain attention to sound. Riss therefore concluded "that the auditory cortex is necessary for localizing the instantaneous position of a sound" (p. 383).

Information from Electrophysiological Studies

Wherever in the brain the tracts from the two ears converge functionally, it should be possible to find interaction between the electrophysiological responses. Kemp and Robinson (1937a) recorded from the brain stem of the anaesthetized cat while presenting monaural or binaural tones or clicks. They interpreted their results as showing no binaural

interaction at the level of the lateral lemniscus, arguing "against the convergence of the tracts from the two ears in the cochlear nuclei or superior olivary complex" (p. 322).

The relative representation of the two ears at the auditory cortex was next investigated in several studies, beginning with that of Bremer and Dow on the cat (1939). Bremer and Dow reported that the response to stimulation of either ear was greater at the contralateral cortex. In contrast to this, Woolsey and Walzl (1942) concluded:

each cochlea is bilaterally represented in the primary projection area, and impulses from corresponding points of each cochlea terminate in common areas of each hemisphere. Ipsilateral and contralateral representations for each cochlea appear to be equal (p. 341).

Tunturi (1944, 1946), working on the dog, reported that each cochlea is represented over the whole extent of both auditory cortices and that each is represented slightly more strongly at the contralateral cortex. Rosenzweig (1951) measured the amplitudes of series of ipsilateral and contralateral cortical responses at 49 electrode placements in five cats. Statistical tests demonstrated that the contralateral response was significantly larger than the ipsilateral at 28 locations; neither was significantly larger at 20 locations, and the ipsilateral response was significantly larger at only 1 location. Thus, the experimenter could tell which ear had been stimulated, either by comparing the amplitudes of responses at the two hemispheres or by inspecting the pattern of amplitudes within a single hemisphere. The stronger representation of each ear in the contralateral hemisphere confirmed the results that Luciani had obtained much earlier in his ablation studies.

Tunturi (1946) had also obtained

evidence of binaural interaction at the cortex. Since his dichotic stimuli were usually separated by intervals of several milliseconds, it remained to determine whether dichotic intervals of a fraction of a millisecond could be preserved in afferent transmission all the way up to the cortex and used there for binaural interaction. Although transmission from the cochlea to the cortex requires about 10 milliseconds, dichotic time intervals of one-tenth of a millisecond were found to affect response amplitudes significantly (Rosenzweig, 1954; Rosenzweig & Rosenblith, 1950). The relation between amplitudes of simultaneous responses at the auditory areas of the two hemispheres was shown to correlate with auditory localization. "At either hemisphere the amplitude of the summated response is larger when the contralateral ear receives the prior stimulus" (Rosenzweig & Rosenblith, 1950, p. 879). "The cortical events were found to parallel in several respects the perceptual phenomena which occur under the same stimulus conditions" (Rosenzweig, 1954, p. 275). Bremer (1952) arrived independently at the hypothesis that the relation between the amplitudes of responses at the two hemispheres is the cerebral index to auditory localization.

Tests were then made for binaural interaction at lower levels of the auditory system. Ades and Brookhart had suggested "that the inferior colliculus with its strong commissural connections and connections to afferent mechanisms may be the principal device responsible for localization" (1950, p. 203). Interaction was found at the inferior colliculus (Coleman, 1953; Rosenzweig & Wyers, 1955), but the importance attributed to the commissural connections by Ades and Brookhart was thrown into doubt by the following observations:

Transecting the commissure of the colliculi does not affect interaction (Rosenzweig & Wyers, 1955); moreover, this transection does not impair auditory localization (Neff & Diamond, 1958). At the colliculi, as at the cortex, recordings made with macroelectrodes show that stimulation of the contralateral ear evokes responses of greater amplitude than does stimulation of the ipsilateral ear (Rosenzweig & Wyers, 1955). Using microelectrodes, Erulkar (1957) found that of 89 single units tested with click stimuli, 23 responded only to stimulation of the contralateral ear, 11 responded only to stimulation of the ipsilateral ear, and 55 responded to stimulation of either ear. For most units that responded to either ear, the latency of response was nevertheless shorter for contralateral than for ipsilateral stimulation. Furthermore, latency showed rather regular changes as the position of the stimulus was moved around the head of the experimental animal (Erulkar, 1959).

Proceeding further down the auditory system, evidence of binaural interaction was also found at the lateral lemniscus (Rosenzweig & Amon, 1955; Rosenzweig & Sutton, 1958). This finding refuted the conclusion of Kemp and Robinson (1937a) that tracts from the two ears do not converge before the colliculi. Kemp and Robinson had found no signs of interaction when they used stimuli dichotic in time, but they gave no details, not even the time intervals employed. Rosenzweig and Sutton, on the contrary, presented measures of the reduction in amplitude of the response to the second stimulus as a function of the dichotic interval.

The lowest level at which binaural interaction occurs may be the superior olivary nuclei. Stotler (1953) reported these findings concerning the

anatomy of the olivary nuclei:

The cells of the medial superior olivary nucleus receive afferent terminals in the form of boutons from both cochlear nuclei. The afferent fibers from the homolateral cochlear nucleus terminate on the lateral pole of the cell, while those from the contralateral cochlear nucleus end in relation to the medial pole. The axons of the medial nucleus enter the homolateral lateral lemniscus (p. 420).

Thus the cells of this nucleus seem ideally situated to integrate information from the two ears. Subsequent electrical recording has in fact shown that some of the superior olivary nuclei show differences in responses depending upon time differences in stimulation of the two ears (Galambos, Schwartzkopff, & Rupert, 1959).

Units in *n. accessorius* proved to be exquisitely sensitive to whether the right ear or the left was stimulated first by paired clicks; the unique physiological and anatomic characteristics of these cells seem relevant to the binocular sound localization problem (p. 527).

Thus the primary neural interactions basic to localization may occur low in the afferent pathways.

It had been suggested in a preliminary report that interaction might occur at as low a level as the cochlea itself, impulses being transmitted from one cochlea to the other (Galambos, Rosenblith, & Rosenzweig, 1950). The transmission appeared to require about 1 millisecond, and some evidence of interaction was obtained when a click at one ear preceded that at the other ear by 1.25 milliseconds. In a later study no evidence of interaural interaction was obtained, using a dichotic interval of 3.6 milliseconds (Rosenblith & Rosenzweig, 1951). It now appears that the latter interval may have been too great, since interaction at the lateral lemniscus can be observed clearly only if the dichotic interval is less than about 3 milliseconds (Rosenzweig & Sutton, 1958). More telling evidence against the occur-

rence of interaural interaction in the cochlea is the failure to find it in the cochlear nucleus under conditions where significant interaction was found in the olivary nucleus and in the lateral lemniscus (Rosenzweig & Amon, 1955). Moreover, as Walsh (1957) has pointed out, interaction at the cochlear level, even if it occurred, could not mediate localization. The time required for transmission from one cochlea to the other seemed to be of the order of 1 millisecond, while localization is obtained with dichotic intervals of one-tenth of a millisecond or less. Thus the lowest level of the auditory system at which binaural interaction has been demonstrated is that of the olivary nuclei. The ablation studies nevertheless suggest that the auditory cortex must be involved if neural interaction in the brain stem is to eventuate in behavioral discrimination of location of sound sources.

Localization with Only One Cerebral Hemisphere?

Walsh (1957) tested 22 patients with various cerebral defects to find how well they could localize sounds. In several of these cases the auditory area of one hemisphere was probably impaired, and in one case the right cerebral hemisphere had been removed. Most of these patients, including the case of hemispherectomy, were able to localize on the basis of time differences between clicks presented at the two ears by means of earphones. With most subjects only a few dichotic intervals were employed, and thresholds were not determined. The intervals were usually of the order of 300 to 500 microseconds—several times the threshold intervals commonly reported for normal subjects. The hemispherectomized patient could localize with intervals of 410 and 190 microseconds but not

with an interval of 125 milliseconds. Walsh concluded, "The sensitivity to binaural time differences is retained after the loss of the auditory cortex on one side" (p. 248). Since thresholds were not determined, it cannot be concluded that there was no impairment of localization with loss of the auditory cortex of one hemisphere.

Accurate thresholds for dichotic stimulation were determined for brain injured subjects in a study reported briefly by Teuber and Diamond (1956). Twenty patients with penetrating brain injuries, 14 of them unilateral, were compared with 10 control subjects who had injuries of peripheral nerves. The brain injured subjects, compared to the controls, required a significantly larger dichotic interval to shift a click from the median plane; the thresholds were 225 and 105 microseconds, respectively. Similarly, the difference in intensities at the two ears necessary to shift a click from the center location was significantly larger for brain injured than for control subjects; the thresholds were 11 and 5 db., respectively. Subjects with unilateral lesions in the right hemisphere required greater intensity on the left than on the right side in order to judge the sound at the midline, and conversely for subjects with left unilateral lesions. (This is similar to the results found by Luciani and Sepilli—1886—with unilateral ablations.) No such directional characteristic was found in the impairment of judgments involving dichotic time. The subjects who were impaired with respect to thresholds for dichotic time were not necessarily impaired with respect to thresholds for dichotic intensity, and conversely. This suggested that the neural mechanisms for localization based on these two cues might not be identical.

Coleman (1959) recorded electrical responses from several positions on the auditory cortex of anaesthetized cats while either moving a click source around the animal's head or varying dichotic time and intensity of clicks produced at the two ears. The relative amplitudes of responses at different electrode positions varied with the location of the sound source or with the dichotic conditions. Some points gave larger responses to contralateral and some to ipsilateral stimuli. "These data suggest that angular location of auditory stimuli may be represented in the auditory cortex of one hemisphere by means of a place principle" (p. 40).

These observations are hard to reconcile with the results of experiments in which electrical stimuli were applied to the auditory area of one cerebral hemisphere (Ferrier, 1890; Penfield & Rasmussen, 1950). It will be remembered that human subjects, under these conditions, usually hear a sound contralaterally to the side stimulated; sometimes they hear the sound on both sides, but never do they hear it ipsilaterally to the stimulation. This suggests that it is not possible to excite points in the right hemisphere that represent locations in space to the right of the listener, nor points in the left hemisphere that represent locations to the left of the listener.

On the basis of his findings Coleman was inclined to reject what he termed "the bilateral ratio theory" of Rosenzweig. No sure conclusion concerning ability to localize can be drawn from Coleman's observations, for the position of a sound source is also represented at the olivary nucleus, yet a cat cannot localize accurately using only the lower brain centers. Should Walsh's conclusion be substantiated that the cortex of one hemisphere suffices for normal

auditory localization, then Coleman's finding may take on increased importance.

Formulation of a comprehensive hypothesis about the mechanisms of binaural localization requires an answer to the question whether the cortex of a single hemisphere is sufficient for normally accurate localization. This question may be resolved by precise determination of the capacity to localize in patients or animals with complete unilateral destruction of the auditory cortex.

SUMMARY AND CONCLUSION

From the work of Venturi in 1800 until about 1920, the perception of location of a sound source was generally considered to be a judgment arrived at by comparing differences in the stimulation at the two ears. While Venturi showed that monaural localization was possible, if the listener could move his head during the presentation of the sound, the chief interest has always been in binaural localization. Over most of this period, dichotic difference in intensity was considered to be the only or the main stimulus basis for localization. Comparison of dichotic intensities seemed to be a plausible explanation for localization, even though Alison pointed out in 1858 that a sound nearer one ear is heard at that side only and seems to be suppressed in the other ear. While the judgmental interpretation and

the intensity hypothesis reigned, there was little incentive and little effort to work out the physiological mechanisms of auditory localization. This was true in spite of noteworthy advances in knowledge of the anatomy of the afferent auditory system, beginning in the 1870s.

The establishment of the dichotic time hypothesis at the end of the first world war was quickly followed by abandonment of the judgmental position. This position could no longer be maintained when it was realized that the time intervals on which localization is based are too small to be perceived as intervals; only a single localized sound is heard. Psychologists soon proposed a number of speculative mechanisms for localization, involving interaction of neural impulses converging from the two ears upon some central locus. During the last 25 years a number of experimenters have brought ablation and electrophysiological techniques to bear on the problems of localization. They have recently shown that the cortex is required for binaural localization, although neural interaction first occurs low in the brain stem. Some evidence suggests that the cortex of a single hemisphere may be sufficient to permit localization. A completely satisfactory hypothesis of the mechanisms of binaural auditory localization, including both cortical and subcortical components, is yet to be presented.

REFERENCES

- ADAMS, H. W., & BROOKHART, J. M. The central auditory pathway. *J. Neurophysiol.*, 1950, 13, 189-205.
- AGGAZZOTTI, A. Sul più piccolo intervallo di tempo percettibile nei processi psichici. *Arch. Fisiol.*, 1911, 9, 523-574.
- ALISON, S. S. On the differential stethophone, and some new phenomena observed by it. *Phil. Mag.*, 1858, 16, 385-395.
- BARD, P., & RIOCH, D. E. A study of four cats deprived of neocortex and additional portions of the forebrain. *Bull. Johns Hopkins Hosp.*, 1937, 60, 73-147.
- BORING, E. G. Auditory theory with special reference to intensity, volume and localization. *Amer. J. Psychol.*, 1926, 37, 157-188.
- BORING, E. G. *Sensation and perception in the history of experimental psychology*. New York: Appleton-Century, 1942.
- BOWLKER, T. J. On the factors serving to de-

- termine the direction of sound. *Phil. Mag.*, 1908, 15, 318-332.
- BREMER, F. Les aires auditives de l'écorce cérébrale. In, *La surdité: Sa mesure et sa correction*. Paris: Maloine, 1952.
- BREMER, F., & DOW, R. S. The acoustic area of the cerebral cortex in the cat. *J. Neurophysiol.*, 1939, 2, 308-318.
- BUNCH, C. C. Auditory acuity after removal of the entire right cerebral hemisphere. *J. Amer. Med. Ass.*, 1928, 90, 2102.
- CERRY, C. *On human communication*. New York: Wiley, 1957.
- COLEMAN, P. D. An electrophysiological study of the interaction between responses to successive clicks in the inferior colliculus of the cat. Unpublished doctoral dissertation, University of Rochester, 1953.
- COLEMAN, P. D. Cortical correlates of auditory localization. *Science*, 1959, 130, 39-40.
- COTZIN, M., & DALLENBACH, K. M. Facial vision: The role of pitch and loudness in the perception of obstacles by the blind. *Amer. J. Psychol.*, 1950, 63, 485-515.
- DOVE, H. W. Eine akustische Interferenz. *Ann. Phys. Chem.*, 1857, 101, 492-494.
- ERULKAR, S. D. The responses of the inferior colliculus of the cat to binaural stimulation. *J. Physiol.*, 1957, 138, 34-35. (Abstract)
- ERULKAR, S. D. The responses of single units of the inferior colliculus of the cat to acoustic stimulation. *Proc. Roy. Soc. Lond., Ser. B*, 1959, 150, 336-355.
- FERRIER, D. *The Croonian lectures on cerebral localization*. London: Smith, Elder, 1890.
- GALAMBOS, R., ROSENBLITH, W. A., & ROSENZWEIG, M. R. Physiological evidence for a cochleo-cochlear pathway in the cat. *Experientia*, 1950, 6, 438-440.
- GALAMBOS, R., SCHWARTZKOPFF, J., & RUPERT, A. Microelectrode study of superior olivary nuclei. *Amer. J. Physiol.*, 1959, 197, 527-536.
- GIRDEN, E. The role of the auditory area of the cortex. *Amer. J. Psychol.*, 1940, 53, 371-383.
- GRIFFIN, D. R. *Listening in the dark*. New Haven: Yale Univ. Press, 1958.
- HESCHL, R. L. *Ueber die vorderquere Schläfenwindung des menschlichen Grosshirns*. Vienna: Braumüller, 1878.
- JEFFRESS, L. A. A place theory of sound localization. *J. comp. physiol. Psychol.*, 1948, 41, 35-39.
- JEFFRESS, L. A. Medial geniculate body: A disavowal. *J. Acoust. Soc. Amer.*, 1958, 30, 802-803.
- JURINE, L. Extrait des expériences de Jurine sur les chauves-souris qu'on a privé de la vue. *J. Phys.*, 1798, 46, 145-148.
- KEMP, E. H., & ROBINSON, E. H. Electric responses of the brain stem to bilateral auditory stimulation. *Amer. J. Physiol.*, 1937, 120, 316-322. (a)
- KEMP, E. H., & ROBINSON, E. H. Electric responses of the brain stem to unilateral auditory stimulation. *Amer. J. Physiol.*, 1937, 120, 304-315. (b)
- KLEMM, O. Sammelreferat über die Lokalisation von Schallreizen. *Kongr. exp. Psychol.*, 1914, Ber. 6, 169-258.
- KLEMM, O. Untersuchungen über die Lokalisation von Schallreizen: III. Mittelung: Ueber den Anteil des beidohrigen Hörens. *Arch. ges. Psychol.*, 1918, 38, 71-114.
- KLEMM, O. Ueber den Einfluss des binauralen Zeitunterschiedes auf die Lokalisation. *Arch. ges. Psychol.*, 1920, 40, 117-146.
- KREIDL, A., & GATSCHER, S. Ueber die dichotische Zeitschwelle. *Pflüg. Arch. ges. Physiol.*, 1923, 200, 366-373.
- LUCIANI, L. On the sensorial localisations in the cortex cerebri. *Brain*, 1884, 7, 145-160.
- LUCIANI, L., & SEPPILLI, G. *Die Funktions-Lokalisation auf der Grosshirnrinde*. (Trans. by M. O. Fraenkel) Leipzig: Denicke's, 1886.
- LUCIANI, L., & TAMBURINI, A. *Ricerche sperimentali sui centri psico-sensori corticali*. Reggio Emilia: Calderini, 1879.
- MAGENDIE, F. *An elementary compendium of physiology*. (4th ed.) (Trans. by E. Milligan) Edinburgh: Carfrae, 1831.
- MALLOCK, A. Note on the sensibility of the ear to the direction of explosive sounds. *Proc. Roy. Soc. Lond., Ser. A*, 1908, 80, 110-112.
- MÜLLER, J. *Handbuch der Physiologie des Menschen*. Vol. 2. Coblenz: Holscher, 1840.
- NEFF, W. D., & DIAMOND, I. T. The neural basis of auditory discrimination. In H. F. Harlow & C. N. Woolsey (Eds.), *Biological and biochemical bases of behavior*. Madison: Univ. Wisconsin Press, 1958.
- NEFF, W. D., FISHER, J. F., DIAMOND, I. T., & YELA, M. Role of auditory cortex in discrimination requiring localization of sound in space. *J. Neurophysiol.*, 1956, 19, 500-512.
- PAVLOV, I. P. *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex*. (Trans. by G. V. Anrep) London: Oxford Univ. Press, 1927.
- PENFIELD, W., & RASMUSSEN, T. *The cerebral cortex of man*. New York: Macmillan, 1950.
- PESTALOZZA, G., & DAVIS, H. Electric responses of the guinea pig ear to high audio frequencies. *Amer. J. Physiol.*, 1956, 185, 595-609.
- PIERCE, A. H. *Studies in auditory and visual space perception*. New York: Longmans, Green, 1901.

- PIÉRON, H. L'orientation auditive latérale. *Année psychol.*, 1922, 23, 186-214.
- RAYLEIGH, Lord. Acoustical observations. *Phil. Mag.*, 1877, 3, 456-464.
- RAYLEIGH, Lord. On our perception of sound direction. *Phil. Mag.*, 1907, 13, 214-232.
- RISS, W. Effect of bilateral temporal cortical ablation on discrimination of sound direction. *J. Neurophysiol.*, 1959, 22, 374-384.
- ROSENBLITH, W. A., & ROSENZWEIG, M. R. Electrical responses to acoustic clicks: Influence of electrode location in cats. *J. Acoust. Soc. Amer.*, 1951, 23, 583-588.
- ROSENZWEIG, M. R. Representations of the two ears at the auditory cortex. *Amer. J. Physiol.*, 1951, 167, 147-158.
- ROSENZWEIG, M. R. Cortical correlates of auditory localization and of related perceptual phenomena. *J. comp. physiol. Psychol.*, 1954, 47, 269-276.
- ROSENZWEIG, M. R., & AMON, A. H. Binaural interaction in the medulla of the cat. *Experientia*, 1955, 11, 498-500.
- ROSENZWEIG, M. R., & ROSENBLITH, W. A. Some electrophysiological correlates of the perception of successive clicks. *J. Acoust. Soc. Amer.*, 1950, 22, 878-880.
- ROSENZWEIG, M. R., & SUTTON, D. Binaural interaction in lateral lemniscus of cat. *J. Neurophysiol.*, 1958, 21, 17-23.
- ROSENZWEIG, M. R., & WYERS, E. J. Binaural interaction at the inferior colliculi. *J. comp. physiol. Psychol.*, 1955, 48, 426-431.
- SPALLANZANI, M. Observations on the organs of vision in bats. *Phil. Mag.*, 1798, 1, 134-136.
- STENGER, —. Simulation und Dissimulation von Ohrkrankheiten und deren Feststellung. *Dtsch. med. Wschr.*, 1907, 33, 970-973.
- STEVENS, S. S., & DAVIS, H. *Hearing: Its psychology and physiology*. New York: Wiley, 1938.
- STOTLER, W. A. An experimental study of the cells and connections of the superior olivary complex of the cat. *J. comp. Neurol.*, 1953, 98, 401-432.
- SUPA, M., COTZIN, M., & DALLENBACH, K. M. "Facial vision": The perception of obstacles by the blind. *Amer. J. Psychol.*, 1944, 57, 133-183.
- TEN CATE, J. Akustische und optische Reaktionen der Katzen nach teilweisen und totalen Extirpationen des Neopallismus. *Arch. Néerl. Physiol.*, 1934, 19, 191-264.
- TEUBER, H. L., & DIAMOND, S. Effects of brain injury on binaural localization of sounds. Paper read at Eastern Psychological Association, Atlantic City, March 1956.
- THOMPSON, S. P. Phenomena of binaural audition. *Phil. Mag.*, 1877, 4, 274-276.
- THOMPSON, S. P. Phenomena of binaural audition. *Phil. Mag.*, 1878, 6, 383-391.
- TRIMBLE, O. C. The theory of sound localization: A restatement. *Psychol. Rev.*, 1928, 35, 515-523.
- TUNTURI, A. R. Audiofrequency localization in the acoustic cortex of the dog. *Amer. J. Physiol.*, 1944, 141, 397-403.
- TUNTURI, A. R. A study of the pathway from the medial geniculate body to the acoustic cortex in the dog. *Amer. J. Physiol.*, 1946, 147, 311-319.
- UPTON, M. Differential sensitivity in sound localization. *Proc. Nat. Acad. Sci., Wash.*, 1936, 22, 409-412.
- VENTURI, G. Riflessioni sulla conoscenza dello spazio che noi possiamo ricavare dall'udito. In, *Indagine fisica sui colori*. Modena: Società Tipografica, 1801.
- VENTURI, J. B. Considerations sur la connaissance de l'étendue que nous donne le sens de l'ouïe. *Mag. encycl. ou J. Lett. Arts* (later title: *Ann. encycl.*), 1796, 3, 29-37.
- VENTURI, J. B. Betrachtungen über die Erkenntnis der Entfernung, die wir durch das Werkzeug des Gehörs erhalten. *Arch. Physiol. (Voigt's Mag.)*, 1800, 5, 383-392. (a)
- VENTURI, J. B. Betrachtungen über die Erkenntnis des Raums, durch den Sinn des Gehörs. *Mag. neu. Zustand Naturkd. (Reil's Arch.)*, 1800, 2, 1-16. (b)
- VON BÉKÉSY, G. Ueber das Richtungshören bei einer Zeitdifferenz oder Lautstärkenungleichheit der beiderseitigen Schalleinwirkungen. *Phys. Z.*, 1930, 31, 824-835, 857-868.
- VON HORNOSTEL, E. M. Das räumliche Hören. In A. Bethe (Ed.), *Handbuch der normalen und pathologischen Physiologie*. Vol. II. Berlin: Springer, 1926.
- VON HORNOSTEL, E. M., & WERTHEIMER, M. Ueber der Wahrnehmung der Schallrichtung. *SB Preuss. Akad. Wiss.*, 1920, 388-396.
- WALSH, E. G. An investigation of sound localization in patients with neurological abnormalities. *Brain*, 1957, 80, 222-250.
- WATSON, L. A., & TOLAN, T. *Hearing tests and hearing instruments*. Baltimore: Williams & Wilkins, 1949.
- WOODWORTH, R. S. *Experimental psychology*. New York: Holt, 1938.
- WOOLSEY, C. N., & WALZL, E. M. Topical projection of nerve fibers from local regions of the cochlea to the cerebral cortex of the cat. *Bull. Johns Hopkins Hosp.*, 1942, 71, 315-344.

(Received April 18, 1960)

PHYSIOLOGICAL EFFECTS OF "HYPNOSIS"¹

THEODORE XENOPHON BARBER

Worcester Foundation for Experimental Biology and Medfield State Hospital, Massachusetts

This paper reviews two series of investigations: one series indicating that sensory, circulatory, gastrointestinal, and cutaneous functions can be altered by means of "hypnosis"; and a second series indicating that similar physiological effects can be produced by symbolic stimulation without "hypnosis."

SENSORY ALTERATIONS INDUCED BY HYPNOTIC STIMULATION²

"Hypnotic Color-Blindness"

To induce "color-blindness" in six "trained" hypnotic subjects (Ss), Erickson (1939) employed a complex procedure which included the following: gradual induction of a "profound somnambulistic hypnotic trance"; slow, gradual induction of "total blindness"; awakening of the S in the "blind" condition so that he would experience distress and anxiety; induction of a second "trance" condition; explanations to the S that vision would be restored but that a certain color or colors would not be

detectable; suggestions of amnesia for the critical color or colors; administration of the Ishihara during suggested (green, red, red-green, and total) color-blindness; administration of the Ishihara without suggested color-blindness in the waking state and in "the simple trance state." The results of this involved experiment (which included 13 separate administrations of the Ishihara to each S) appeared to be as follows: all Ss had normal color vision during the waking state and in "the simple trance state"; during suggested color-blindness, the numerals on some of the Ishihara cards were read in the manner characteristic of the green, red, red-green, or total color-blind. Erickson concluded that the hypnotic procedure was effective in inducing "consistent deficiencies in color vision comparable in degree and character with those found in actual color blindness." However, Grether (1940) criticized this conclusion noting that (a) "red-green color-blindness" does not exist in nature (this is a generic term referring to symptoms common to red-blindness and green-blindness); and (b) the deficiencies in color vision found among persons with actual red-blindness, green-blindness, or total color-blindness are "quite different" from those which Erickson attempted to induce. Harriman (1942) repeated part of Erickson's procedure, suggesting amnesia for red and then for green to 10 "deeply hypnotized" Ss; although these Ss responded to the Ishihara in a manner similar to Erickson's Ss, Harriman concluded, in accordance

¹ The author is indebted to Louis B. Glass, Alejandro D. Paniagua, David R. Evans, and Harry Freeman for critically reading the manuscript. A condensed version of the paper was presented at the University of Kansas seventh Annual Institute of Research in Clinical Psychology, May 10, 1960.

Research reported by the author and his associates was supported in part by a grant from the National Association for Mental Health and in part by Research Grant MY3253, from the National Institute of Mental Health, United States Public Health Service.

² Since experimental and clinical studies of "hypnotic analgesia" have been recently reviewed elsewhere (Barber, 1959, 1960a), this phenomenon is not included in the following discussion.

with Grether's critique, that the alterations induced "resemble attitudinal changes more closely than they resemble profound changes in sensory content." However, no attempt was made to determine if the lengthy and involved hypnotic procedure employed in the investigation was actually necessary to induce such "visual anomalies."

Barber and Deeley (1961) hypothesized that normal Ss, instructed to remain inattentive to red or green, give similar responses to the Ishihara as "hypnotic color-blind" Ss. As a preliminary test of color vision, the American Optical Company Pseudo-Isochromatic Plates were administered to 10 normal Ss. The S was then presented with the Ishihara plates and instructed as follows: "Now look at these cards. As I present each card, try as hard as you possibly can to pay no attention to the red. Look carefully at the rest of the card, but ignore the red; just don't let yourself see it." After completing this task the Ishihara cards were presented again and similar instructions were given to "try as hard as you possibly can to pay no attention to the green." Finally, the S was instructed to report what he naturally saw on the Ishihara plates. The results were as follows: (a) The responses to the Pseudo-Isochromatic Plates and to the final administration of the Ishihara indicated normal color vision in all Ss. (b) When instructed to "pay no attention" to red and then to green, 92 of 320 (28.8%) of the total responses of the 10 normal Ss were similar to the responses expected from persons with natural red-blindness or green-blindness. Of the 320 responses given to the Ishihara by Harriman's 10 "deeply hypnotized" Ss during suggested red-blindness and green-blind-

ness, 85 (or 26.6%) were similar to the responses expected from the red-blind or green-blind. In brief, this experiment appears to indicate that normal persons who have been instructed to concentrate away from red or green give similar responses on the Ishihara as "deeply hypnotized" Ss who have been given elaborate suggestions to induce color-blindness.

"Hypnotic Blindness"

Are hypnotic suggestions of total blindness effective in altering physiological processes related to vision? Hernández-Pedón and Donoso (1959) recently published a neurophysiological experiment which, although not a direct study of hypnotically-induced blindness, nevertheless promises to contribute to our understanding of this phenomenon. Electrodes were deeply implanted in the occipital lobes of five patients who had undergone trephination for diagnostic explorations. With the occipital electrodes in place, the room was darkened and the patient was stimulated by electronic lamp flashes at the rate of 1/millisecond. In each case the electrographic recordings showed an evoked potential simultaneous with the photic stimulation. Subsequently, when two of the patients, whom the experimenters judged to be especially "suggestible," were given repeated verbal suggestions that the light intensity was greater than that actually applied, the electrographic recordings indicated an enhancement of the photically evoked potentials; when given the suggestion that the intensity of the light had diminished, while it actually remained constant, the recordings showed a diminution of the evoked potentials. However, in related experiments the same investi-

gators demonstrated that the magnitude of the photically evoked potentials was consistently reduced whenever "the attention of the subject was distracted," e.g., when instructed to solve a difficult arithmetic problem mentally or when asked to recall an interesting experience. From these experiments and from a series of related studies by other workers summarized in the paper, the authors suggest that during "voluntary attention" as well as by suggestion, transmission of photic impulses is modified at the retina by centrifugal influences. These influences, acting during wakefulness, are probably related to organized activity of the reticular formation of the brain stem under the control of the cortex (p. 394).

In earlier studies, Dorcus (1937), Lundholm and Lowenbach (1942), and other workers had noted that the pupillary reaction to light stimulation is not altered during "hypnotic blindness." However, since pupillary constriction to light is found during some types of organic blindness (e.g., bilateral destruction of the occipital visual areas—Madow, 1958), this response is not a satisfactory index of blindness and workers in this area have generally focused on an ostensibly more satisfactory response—alpha blocking on the electroencephalogram (EEG).

Alpha blocking to photic stimulation appears to be a totally involuntary response which is almost always present in normal persons and never present in the blind. A series of investigations has demonstrated that (a) when the room is darkened and the eyes are closed, most normal persons typically show an alpha rhythm on the EEG (consisting of waves with a frequency of 8 to 13 cycles per second and an amplitude of about 50 microvolts); (b) a light flashed into the closed eyes of these

individuals is almost always effective in causing "alpha block" or "alpha desynchronization" (i.e., in replacing the alpha rhythm with small fast waves) within 0.4 second (Jasper & Carmichael, 1935); and (c) persons with total blindness of neurological origin do not show alpha blocking under these conditions (Callahan & Redlich, 1946).

Lundholm and Lowenbach (1942), Barker and Burgwin (1948), and Ford and Yeager (1948) found that hypnotic suggestions of blindness did *not* prevent alpha blocking when the Ss opened their eyes in an illuminated room. However, these experiments are based on a methodological error: In normal persons the act of opening the eyes *per se*—whether in darkness or in an illuminated room—almost invariably results in alpha desynchronization (Loomis, Harvey, & Hobart, 1936; Yeager & Larsen, 1957). To determine if hypnotic suggestions of blindness are effective in preventing the alpha desynchronization which normally occurs after visual stimulation, it is therefore necessary for the S either to keep his eyes continuously open or continuously closed during the experiment. These conditions have been met in three investigations. Loomis et al. (1936) demonstrated that when total blindness was suggested to an excellent hypnotic S whose eyes were kept open continuously with adhesive tape, the alpha rhythm did *not* show desynchronization during photic stimulation. This was repeated 16 times with the same results; whether the room was illuminated or darkened made no difference whatsoever—the alpha rhythm was continuously present until the S was told that he could once again see. In a subsequent experiment, Schwarz, Bickford, and Rasmussen (1955) found that after

suggestions of blindness 7 of 11 hypnotic Ss (with eyes taped open) showed occasional alpha waves when the room was illuminated. In a more recent study, Yeager and Larsen (1957) instructed five Ss to keep their eyes continuously closed during the experiment. Hypnotic and post-hypnotic suggestions were given that the S would not be aware of the light stimulation. In the majority of trials, no alpha blocking occurred when light fell upon the closed eyes.

The above studies indicate that hypnotic suggestions of blindness are at times effective in eliminating an involuntary physiological response which normally follows visual stimulation, viz., alpha blocking on the electroencephalogram. However, a similar effect can be demonstrated in Ss who have not been given an "hypnotic induction" and who do not appear to be in "the trance state." Loomis et al. (1936) found that when a uniformly illuminated bowl was placed over the eyes of a normal person who was instructed not to focus on any specific part of the light pattern, the alpha waves appeared fairly regularly. Gerard (1951) writes:

With a little practice I can look directly at a 100-watt light . . . and, by deliberately paying no attention to it, I can have my alpha waves remain perfectly intact; then with no change except what I can describe in no other way than as directing my attention to the light, have them immediately disappear (p. 94). (Quoted by permission of John Wiley & Sons.)

Jasper and Cruikshank (1937) have published similar findings. In brief, although some "hypnotized" Ss, who have been given suggestions of blindness, continue to show an occipital alpha rhythm part of the time or all of the time when stimulated by light, a similar effect can be demonstrated

in normal persons who are instructed to "pay no attention" to visual stimuli.

In a recent study Schwarz et al. (1955) found that five "hypnotized" Ss who had been given suggestions of blindness did not show eye movements when urged to look at an object. The restriction of eye movements was indicated both by electromyographic eye leads and by the marked suppression of lambda waves on the EEG. These investigators suggest that the restriction of eye movements during hypnotic blindness "is an attempt to shut off all alerting stimuli that might interfere with the successful accomplishment of the suggestion." Along similar lines, Barber (1958b) presented evidence indicating that seven somnambulist hypnotic Ss deliberately refused to look at an object which they had been told that they could not see; observation of eye movements indicated that they typically focused on all parts of the room *except* where the object was situated. When interviewed after the experiment, most of the Ss readily admitted that they purposely refused to carry out the active process of turning the head and focusing the eyes on the object, e.g., "I was almost carefully not looking at it," "I kept looking around it or not on it."

In an earlier study, Pattie (1935) gave five good hypnotic Ss the suggestion that they were blind in one eye. Four responded to a series of visual tests (stereoscope, perimetry, filters, Flees' box, plotting the blind spot, ophthalmological examination) with normal vision in both eyes; however, one S responded to all tests as if she were actually blind in one eye. In a second experiment the "blind" S was given a more complicated filter test; the results indicated that the

"blind" eye was not impaired to the slightest degree and Pattie concluded that the "former tests were thus invalidated." When questioned in a subsequent hypnotic session, the *S* revealed, after much resistance, that she had given a convincing demonstration of uniocular blindness because of the following: during the stereoscopic test the two images were separated a second after exposure and this gave her the necessary knowledge to fake the test; she had practiced determining the blind spot at home after the experimenter had first attempted to plot it; on the Flees' box with crossed images she "saw there were mirrors in there and figured somehow that the one on the left was supposed to be seen with the right eye," etc.

The above studies appear to indicate that the "good" hypnotic *S*, who has been given suggestions of blindness; purposely attempts to inhibit responses to visual stimuli. This suggests the following hypothesis which can be easily confirmed or disproved: The responses to photic stimulation which characterize "deeply hypnotized" *Ss* who have been given suggestions of blindness can be duplicated by normal persons who are asked to remain inattentive and unresponsive to visual stimuli.

"Hypnotic Deafness"

Can significant alterations in auditory functions be demonstrated in the hypnotized person following suggestions of deafness? Fisher (1932) and Erickson (1938b) approached this question by investigating the effect of hypnotically-induced "deafness" on conditioned responses to acoustic stimuli. Fisher found that during posthypnotic deafness one *S* did not show a patellar response which had been conditioned to the sound of a

bell; Erickson similarly demonstrated that after hypnotic suggestions of deafness two *Ss* failed to show a hand-withdrawal response conditioned to the sound of a buzzer. Although both investigators interpret the failure to show conditioned responses to auditory stimuli as a sign of deafness, earlier experiments, reviewed by Hilgard and Marquis (1940, p. 35, pp. 269-270), which indicate that such conditioned responses can be voluntarily inhibited, suggest a second interpretation, namely, that the "hypnotic deaf" *Ss* perceived the sound stimulus but purposely inhibited the response. Some support for this interpretation is offered by the kymographic tracings reproduced in Fisher's paper which show an aborted patellar response to some of the sound stimuli. Additional evidence is presented by Lundholm (1928) who, like Erickson, conditioned a hand-withdrawal response to an auditory stimulus; although the *S* in this case did not show the conditioned response after hypnotic suggestions of deafness, he later admitted "having heard the click, having felt an impulse to withdraw on click without shock, and having resisted and inhibited that impulse" (p. 340).

As an additional index of deafness, Erickson (1938a) noted that his *Ss* did not show startle responses to sudden loud sounds. Other investigations, however, again suggest the possibility that the *Ss* may have perceived the sound and purposely inhibited the startle response; for instance, Dynes (1932) reported that three "hypnotic deaf" *Ss*, who did not show overt startle responses when a pistol was unexpectedly fired, admitted after the experiment that they heard the sound, and Kline, Guze, and Haggerty (1954) demon-

strated that a "hypnotic deaf" *S* who failed to show both conditioned responses to auditory stimuli and startle reflexes to sudden loud sounds showed clear-cut responses to auditory stimuli when tested by a method employing delayed speech feedback.

The latter experiment merits further comment. In the normal person, feeding back his speech through tape recording amplification and earphones with a delay of one-quarter second has been reported to produce an impairment in subsequent speech. Most commonly this speech disturbance involves stammering, stuttering, perseveration, and marked loss in speed and tempo. Kline et al. (1954) found that such delayed speech feedback produced distinct impairment in speech performance in an excellent hypnotic *S* who had been given suggestions of deafness. However, as compared with his "waking" performance, the *S* showed less slurring, stuttering, and stammering, appeared more calm, and did not show discomfort. The investigators concluded that the hypnotic suggestions of deafness were effective in inducing a "set," or in "gearing" the *S*, "to give minimal response to the excruciating intensity and the constant interference of the feed-back of his own voice" without in any way inducing "deafness in the usual sense." However, no attempt was made to determine if the *S* would have shown a similar ability to tolerate the speech-disturbing stimulation during the "waking" experiment if he had been carefully instructed and motivated to remain inattentive to or to "concentrate away from" the stimulation. Further experiments are required to determine if normal persons are able to duplicate the behavior of this "deeply hypnotized" *S* when instructed in this manner.

Malmo, Boag, and Raginsky (1954) have reported comparable findings. After appropriate suggestions to induce deafness, two somnambulistic *Ss* denied auditory sensations and showed significantly reduced motor reactions to sudden auditory stimulation; however, myographic recordings from eye muscles showed a strong blink reaction in both *Ss* at each presentation of the auditory stimulus. Sternomastoid tracings indicated that one *S* showed slight startle responses to all stimuli and the other *S* showed a strong startle reaction to the first presentation of the stimulus and slight startle reactions to subsequent stimuli. Other data presented in the report (e.g., introspective reports and myographic tracings indicating a higher level of tension in the chin muscles under hypnosis as compared to the control condition) permit the following interpretation of the findings: (a) the *Ss* were unable to inhibit blink responses to the auditory stimuli; (b) since the first presentation of the auditory stimulus was more or less unexpected, one *S* failed to inhibit the startle response; (c) since the second and subsequent stimuli were expected, both *Ss* were able, to a great extent, to inhibit startle responses. In an earlier study Malmo and his collaborators (Malmo, Davis, & Barza, 1952) found that, when unexpectedly presented with an intense auditory stimulus, a hysterical "deaf" patient also showed a gross startle response on the myograph; a control case of middle-ear deafness, studied by the same techniques, showed no blink reaction and no startle response to any presentation of the auditory stimulus.

In an earlier study Pattie (1950) gave four somnambulistic hypnotic *Ss* suggestions of unilateral deafness.

The Ss appeared to accept the suggestions, insisting that they could not hear in one ear. However, when auditory stimuli were presented in such a manner that they could not determine which ear was being stimulated, they showed normal hearing in both ears.

The above findings—that "hypnotic deaf" Ss purposely inhibit conditioned responses to auditory stimuli (Lundholm), appear to inhibit startle responses to sudden acoustic stimuli (Malmo et al., 1952), show a calmer attitude and less tension during speech-disturbing auditory stimulation but no sign of actual deafness (Kline et al., 1954), and do not show "deafness" in one ear when unable to determine which ear is being stimulated (Pattie, 1950)—suggest a similar hypothesis as the studies of "hypnotic blindness" reviewed in the preceding section of this paper: if carefully instructed and motivated to "concentrate away from" auditory stimulation, normal persons show similar responses to acoustic stimuli as "hypnotic deaf" Ss.

THE EFFECT OF HYPNOTIC STIMULATION ON CIRCULATORY FUNCTIONS

Effect of Hypnotic Stimulation on Vasomotor Functions

The evidence at present indicates that localized vasoconstriction and vasodilation (and a concomitant localized skin temperature alteration) can be induced in some hypnotized persons by appropriate verbal stimulation. McDowell (1959) found that a good hypnotic S showed erythema with vasodilation and increase in skin temperature of the right leg following suggestions that the leg was immersed in warm water. In a careful experiment, Chapman, Goodell, and Wolff (1959) suggested to 13 Ss

"as soon as a state of moderate to deep hypnosis had been established," that one arm was either "normal" or that it was numb, wooden, and devoid of sensation ("anesthetic"). The arm was then exposed on three spots, blackened with India ink, to a standard thermal stimulus (500 millicalories/second/centimeter² for 3 seconds). After an interval of 15 to 30 minutes "during which [time] hypnosis was continued," it was suggested that the other arm was tender, painful, burning, damaged, and exceedingly sensitive ("vulnerable") and the same standard noxious stimulation was applied. The results of 40 experiments with the 13 Ss were as follows: In 30 experiments the inflammatory reaction and tissue damage following the noxious stimulation was greater in the "vulnerable" arm, in 2 experiments the reaction was greater in the "anesthetic" arm, and in 8 experiments no difference was noted. Plethysmographic and skin temperature recordings indicated that following the noxious stimulation local vasodilation and elevation in skin temperature was larger in magnitude and persisted longer in the "vulnerable" arm. This experiment should be repeated with un hypnotized Ss who are instructed to *imagine* one arm as "devoid of sensation" and the other arm as "exceedingly sensitive." The data summarized below suggest that at least some of the effects reported in this study—localized vasodilation and elevation in skin temperature—can be induced by symbolic stimulation in some individuals who do not appear to be "in a state of moderate to deep hypnosis."

When attempting to condition local vasoconstriction and vasodilation to verbal stimuli, Menzies (1941) found that the conditioning procedure could

be dispensed with in some cases; some persons, who had *not* participated in the experimental conditioning, showed vasodilation in a limb when recalling previous experiences involving warmth of the limb and local vasoconstriction when recalling experiences involving cold. In an earlier study, Hadfield (1920) found that localized changes in skin temperature could be induced by suggestions given to a person "in the waking state." In this case, the *S* had exercised vigorously before the experiment and the temperature of both hands, as measured with the bulb of the thermometer held firmly in the palm, had reached 95°F. Without a preliminary hypnotic procedure, it was suggested that the right arm was becoming cold. Within half an hour the temperature of the right palm fell to 68° while the temperature of the left palm remained at 94°. When subsequently given the suggestion that the right hand was becoming warm, the temperature of the hand rose within 20 minutes to 94°. Although this *S* had previously participated in hypnotic experiments, Hadfield insists that he did not "hypnotize" him during this experiment and that the temperature alterations occurred when the *S* was "entirely in the waking condition."

Cardiac Acceleration Produced by Hypnotic Stimulation

A number of experiments, reviewed by Gorton (1949) and Weitzenhoffer (1953), demonstrate that cardiac acceleration can be produced by hypnotic suggestions which activate the *S* and that cardiac deceleration can be produced by hypnotic suggestions of relaxation, drowsiness, and sleep; however, this finding indicates no more and no less than that an alteration in the "level of arousal" or

"level of activation" (Duffy, 1957; Woodworth & Schlosberg, 1954)—whether induced by stimuli present during various ongoing life situations or induced by symbolic stimulation during a hypnotic experiment—is correlated with an alteration in the heart rate. A more significant question is: Can the heart rate be accelerated or depressed by direct suggestions of such an effect without simultaneously inducing anxiety, emotion, or arousal? Solovey and Milechnin (1957) demonstrated an accelerated pulse rate in 2 out of 23 hypnotic *Ss* following the direct suggestion "Your heart is beating more rapidly." However, the possibility is not excluded that the cardiac acceleration in the two cases was due to emotion or anxiety evoked by the suggestions; one *S* later reported that, when given the suggestion, he imagined himself looking down from a height and feeling someone pushing him on the shoulder and the other *S* stated that, when given the suggestion, he had a feeling of distress. Since relatively large changes in cardiac rate can be demonstrated during alterations in the rate and depth of respiration (Huttenlocher & Westcott, 1957), it also appears plausible that the altered pulse rate in these cases may have been an indirect effect of a change in respiratory pattern.

To demonstrate a *direct* effect of symbolic stimulation on heart rate it is necessary to control at least two factors, "level of arousal" and respiratory rate. To the writer's knowledge only one hypnotic experiment has been published which ostensibly satisfies these criteria: Van Pelt (1954) reported that a somnambulistic hypnotic *S* showed an accelerated cardiac rate following direct suggestions of such an effect without at the

same time showing an altered respiratory rate or emotional arousal. After an "hypnotic induction" procedure, this investigator spoke to the *S* in a quiet tone as follows: "Your heart is beginning to beat faster. It is getting faster and faster. You are perfectly calm, but your heart is beating faster and faster." The electrocardiogram (EKG) showed that the heart rate increased immediately from 78 to 135 beats per minute. Although Van Pelt states that he did not observe a change in the depth and rate of respiration during the tachycardia, it appears possible that an altered respiratory pattern could have been demonstrated if a pneumograph had been employed in the study. However, during the acceleration the *S* appeared calm and the EKG tracing did not show somatic tremors which are typical of nervousness and fear. In a second experiment, in which the same *S* showed cardiac acceleration following suggestions intended to arouse fear—"You are driving a car at a tremendous speed and are heading toward a second car and are going to crash"—the EKG recording showed clear evidence of somatic tremors.

The above study lacks a crucial control; no attempt was made to determine if the *S* could voluntarily accelerate the heart without "hypnosis." Since other workers employing similar procedures with equally "good" hypnotic *Ss* have failed to demonstrate cardiac acceleration (e.g., Jenness & Wible, 1937, failed in 30 attempts with eight somnambulists), it appears plausible that the hypnotic procedure was not a necessary factor in producing this effect. Supporting evidence for this supposition is presented in a series of studies (ca. 20) which demonstrate that some apparently normal persons are able to accelerate the heart volun-

tarily (King, 1920). In most of the reported cases the voluntary tachycardia was on the order of 15 to 40 beats per minute; however, in some cases (Favell & White, 1917; Tarchanoff, 1885) the acceleration was as high as 63 or 75 beats per minute. In all cases the *Ss* denied that they induced the tachycardia by visualizing emotion inducing situations and insisted that they produced the effect by voluntary effort. Some *Ss* showed changes in respiratory pattern during the voluntary tachycardia but in these cases the respiratory alterations varied and could not be correlated with the change in heart rate (Koehler, 1914; Pease, 1889; Tarchanoff, 1885; Van de Velde, 1897); other *Ss* could as readily induce the voluntary acceleration when breathing more or less normally as when showing changes in respiratory pattern (King, 1920; Taylor & Cameron, 1922); and some *Ss* showed no significant change in respiratory pattern on the pneumograph when inducing cardiac acceleration on the order of 40 beats per minute (Favell & White, 1917).

Voluntary acceleration of the heart may not be as uncommon as is generally assumed: Van de Velde found four cases and Tarchanoff five cases when confining their search to relatively small groups of individuals; a number of medical students discovered that they possessed this ability in physiology classes when they attempted to determine the validity of the lecturer's assertion that voluntary cardiac acceleration is not impossible (Ogden & Shock, 1939; West & Savage, 1918).

Cardiac Standstill Induced by "Hypnosis"

Raginsky (1959) demonstrated that hypnotic suggestions are effective in producing cardiac block for a

brief period in an appropriately predisposed person. The *S* in this case was a hospitalized patient whose carotid sinuses had been surgically removed because of periodic fainting episodes with cardiac arrest (Adams-Stokes disease). After the patient "went into a medium to deep hypnotic state," he was instructed "in a tone of considerable urgency to visualize with all clarity possible his worst attack of faintness." The patient "turned pale, limp, and a cold perspiration appeared on his forehead. His pulse was unobtainable . . ." The EKG tracing showed complete auricular and ventricular standstill for a time interval of four beats, followed by a normal sinus-auricular beat. After a rest period of 10 minutes, the experiment was repeated with comparable results. However, no attempt was made to determine if cardiac arrest could be induced in this patient by asking him to visualize his worst attack of faintness *without* a preceding "hypnotic induction" procedure. The case summarized below suggests that the "hypnotic induction" and the "medium to deep hypnotic state" may have been unnecessary in producing this effect.

McClure (1959) found that an appropriately predisposed person could voluntarily produce cardiac standstill. The *S* in this case, a 44-year-old airplane mechanic, had discovered that he could induce a progressive slowing of the pulse by relaxing completely. When asked to induce a diminution of the heart rate in the laboratory, the *S* lay very quietly and allowed respiration to become extremely shallow. The EKG showed sinus arrest for a period of 5 seconds. An EKG tracing taken 1 hour after the experiment was within normal limits. Since the *S* had rheumatic fever at age 7, McClure suggests the

following tentative explanation of this performance:

The underlying cardiac change is believed to be well compensated rheumatic heart disease with aortic valvulitis. The bradycardia and cardiac arrest are probably manifestations of exaggerated vagotonia, induced through some mechanism which, although under voluntary control, is not known to the patient himself. Careful observation did not reveal any breath-holding or Valsalva maneuver in connection with the cessation of heartbeat. Apparently the patient simply abolished all sympathetic tone by complete mental and physical relaxation (pp. 440-441). (Quoted by permission of *California Medicine*.)

Electrocardiographic Alterations Induced by Hypnotic Stimulation

Bennett and Scott (1949) found that one of five excellent hypnotic *Ss* showed tachycardia and T wave abnormalities on the EKG—lowering or disappearance of T in Leads I, II, and III—within 2 minutes following suggestions intended to induce anxiety and anger. The *S* in this case was an emotionally stable and well-adjusted young male with no history of cardiac disorders and with an otherwise normal EKG. Since such EKG abnormalities are not normally associated with tachycardia, two electrocardiographers, who were not informed of the experimental conditions under which the tracings were made, interpreted the records as indicating coronary artery disease or acute rheumatic fever. Finding in a subsequent study with the same *S* that subcutaneous administration of epinephrine elicited lower T waves in Leads I and II than those found during the control experiment, the authors suggest that the EKG alterations induced during the hypnotic experiment may have been an indirect result of sympathetic stimulation and release of epinephrine from the adrenal medulla. Berman, Simonson, and Heron (1954) confirmed this study; employing 14 susceptible

hypnotic Ss with otherwise normal EKG, they found that during hypnotically-induced fear and anxiety, two showed elevation and five showed depression or inversion of T waves. In a second experiment these workers found that although "deep hypnosis" could not be induced in 11 patients with coronary sclerosis and angina pectoris, four showed T wave changes when given emotion inducing suggestions.

The above experiments demonstrate that EKG alterations resembling those found in grave cardiac disorders can be induced in some "hypnotized" Ss by suggestions which evoke fear, anger, or anxiety; however, similar EKG abnormalities have been demonstrated in some normal persons during emotional arousal. Mainzer and Krause (1940) compared the EKG tracings of 53 unselected surgical patients recorded the day before surgery, and on the operating table immediately before the induction of general anesthesia. As compared with the earlier tracings, 40% of the tracings recorded immediately before surgery showed various abnormalities such as S-T depression with T low, inverted, or absent. Along similar lines, Landis and Slight (1929) and Loftus, Gold, and Diethelm (1945) demonstrated that some persons with otherwise normal EKG show abnormalities of the ST segment and the T wave during startle or anxiety; Crede, Chivers, and Shapiro (1951) found that in rare cases mere anticipation of the EKG test is sufficient to produce inverted T waves in normal individuals; and Ljung (1949) published a study of 14 Ss with no evidence of cardiac disease who showed abnormal T waves during apparently slight emotional stimulation. After summarizing these and related in-

vestigations, Weiss (1956) suggests that such EKG effects are found during emotional stimulation in persons who are prone to show an elevation of sympathetic tone and an increase in cardiac metabolism without a corresponding increase in the coronary circulation.

In brief, the above studies on cardiac functions indicate the following:

1. In very rare cases, it is possible to produce cardiac acceleration or complete stoppage of the heart action by appropriate hypnotic stimulation; however, in very rare cases, similar effects can be voluntarily produced by un hypnotized persons.

2. Although some hypnotized persons show EKG alterations resembling those found in organic heart disease following suggestions designed to induce fear, anxiety, or anger, some persons who have not been given an "hypnotic induction" and who do not appear to be "in trance" show similar EKG alterations during emotional stimulation.

EFFECT OF HYPNOTIC STIMULATION ON METABOLIC AND GASTROINTESTINAL FUNCTIONS

Effect of Hypnotic Stimulation on Blood Glucose Levels

A number of experiments appear to indicate that hypnotized persons show an elevation of blood glucose levels when given the direct suggestion that blood sugar will increase. Before discussing these studies, it is appropriate to note the following:

1. The level of blood glucose appears to be closely related to the level of "arousal"; blood sugar tends to increase during anxiety, emotion, or maintained activity and to decrease during relaxation, depression, or sleep (Dunbar, 1954, Ch. 8).

2. The blood glucose level is excessively labile in diabetics, i.e., as

compared with normal persons, diabetics tend to show more extreme alterations in blood sugar content during periods of high or low "arousal" (Hinkle & Wolf, 1953; Mirsky, 1948).

The above postulates suggest that in diabetic patients any procedure (hypnotic or nonhypnotic) which induces relaxation or minimizes excitability will tend to depress the blood sugar level and minimize glycosuria and any procedure which induces arousal or excitability will tend to elevate the blood glucose level and increase glycosuria. Data supporting this hypothesis have been presented by Bauch (1935) in a study of the effects of training in relaxation (Schultz's "autogene training") on seven diabetic patients. Each patient showed a significant decrease in blood sugar levels after becoming proficient in inducing relaxation—insulin dosage was reduced in each case by 10 to 20 units. Apparently, healthy persons do *not* show the same degree of reduction in blood glucose content after achieving the same success in producing relaxation (Schultz & Luthe, 1959). Along similar lines, Mohr (1925) relieved a "full-pledged diabetic" of glycosuria by hypnotic suggestions which were effective in mitigating his "affective excitability" toward certain significant persons in his surroundings and was able to reinstate the glycosuria by suggesting that he would again be upset by these people. This experiment was repeated four times with the same results.

With the above findings in mind, the results reported in two hypnotic experiments become less mysterious. Gigon, Aigner, and Brauch (1926) found that blood sugar tended to be reduced in four hypnotized diabetic patients following suggestions of re-

laxation and suggestions that "the pancreas would secrete insulin and that blood and urine sugar would markedly decrease." Although the reduction in blood glucose in these cases may have been due to the suggestion that "the pancreas would secrete insulin," it appears equally plausible that it was a secondary effect of the suggestions of relaxation. Along similar lines, Stein (quoted by Dunbar, 1954, p. 291) reported that direct suggestions that blood sugar would decrease given to six hypnotized diabetic patients resulted in reduced blood sugar in 47 out of 56 attempts. Again, it appears plausible that the reduced blood glucose in these cases was an indirect result of the suggestions of relaxation given during the "hypnotic induction" procedure. Supporting evidence for this supposition is presented in a second experiment by the same investigator; although Stein used only one diabetic patient in this study, he found that an "hypnotic induction" (apparently consisting of suggestions of quietude, relaxation, and drowsiness) resulted in a significant fall in blood glucose content *without* suggesting that the blood sugar would fall.

Is it possible to elevate the blood sugar level by suggesting to a non-diabetic hypnotic *S* that he is ingesting sugar? Marcus and Sahlgren (1925) found no rise in blood glucose content when four "deeply hypnotized" nondiabetics were given a saccharin solution which they were told was a sugar solution. Similarly, Nielsen and Geert-Jorgensen (1928) found no elevation in the fasting blood sugar level when six excellent hypnotic *Ss* (nondiabetics) were given the suggestion that a glass of water contained large amounts of sugar. In contradistinction to the above, Povorinskij and Finne (1930) found

an elevated blood sugar content in two somnambulistic hypnotic *Ss* after inducing an hallucination of ingesting sugar and honey; however, an elevation in blood glucose could be demonstrated in one of these *Ss* following similar suggestions given during "the waking state." The data presented in the report do not exclude the possibility that the hypnotic suggestions which induced an "hallucination" of ingesting sugar and honey served to "arouse" the subjects or to induce emotional excitement.

Effect of Hypnotic Stimulation on Gastric Functions

The evidence indicates that stomach secretions, hunger contractions, and various other gastrointestinal functions can be influenced by appropriate suggestions given to a hypnotized person. Ikemi (1959) demonstrated that suggestions given during hypnosis of eating a delicious meal resulted in an increase in free acid, total acidity, and quantity of gastric secretions in 34 out of 36 healthy young persons. In an earlier experiment, Heyer (1925) introduced a tube into the stomach of a "deeply hypnotized" *S* and removed the contents. If no secretion occurred within 10 minutes, the *S* was given the suggestion that he was ingesting either meat broth, bread, or milk and the gastric secretions were collected at 5-minute intervals and examined for quantity, acidity, and proteolytic activity. Each of the suggested meals evoked a secretion of approximately 6 to 10 cubic centimeters of "gastric juice" within 10 to 15 minutes and the acidity and proteolytic activity appeared to vary with each food suggested. Delhougne and Hansen (1927) reported a similar study with one somnambulistic *S*. After the *S* was placed in "deep hypnosis," the

stomach and duodenal secretions were aspirated by means of a Reh-fuss tube. Following this, the *S* was given the suggestion that he was ingesting a meal which was rich in protein (Schnitzel), rich in fat (a biscuit thickly covered with butter), or rich in carbohydrate (chocolate and marchpane). Each of the suggested meals evoked secretions of acid, pepsin, and lipase from the stomach and of trypsin, lipase, and diastase from the pancreas. Although the authors do not analyze the data statistically, they conclude that the hallucinated meals were as effective as actual meals in eliciting *specific* secretions from the stomach and pancreas, e.g., the hallucinated protein meal supposedly induced a specific increase in the secretion of pepsin and trypsin, the hallucinated fatty meal supposedly induced a specific increase in the secretion of lipase. This startling conclusion, however, appears to be erroneous; a statistical analysis indicates that the quantity of each of the enzymes found after the three hallucinated meals was not significantly different.

The above studies do not answer a crucial question: Was the "hypnotic induction" and the appearance of "deep trance" on the part of the *Ss* necessary to produce these effects? If the *Ss* had been asked to vividly imagine or to think about eating certain foods (without an "hypnotic induction") would they have shown similar pancreatic and gastric secretory activity? That such may have been the case is suggested by an earlier experiment reported by Luckhardt and Johnston (1924). These investigators also found that when a hypnotized *S* was given suggestions of eating a fictitious meal, he showed an increase in the volume and acidity of the digestive secretions compa-

rable to that found when actually eating a meal; however, in the control experiment, when the investigators merely talked to the *S* about an appetizing meal, he showed similar gastric secretory activity. This finding is not unusual. Miller, Bergeim, Rehfuess, and Hawk (1920) reported that the sound and thought alone of a frying steak gave rise to gastric secretory activity in some normal *Ss*. Employing a *S* with a gastric fistula, Wolf and Wolff (1947) demonstrated that during the "mere discussion" of eating a certain food the output of hydrochloric acid from the parietal cells was essentially the same as when actually ingesting this food. Similar effects have been demonstrated in other parts of the gastrointestinal tract. Bykov (1957) found that in patients with a gall bladder fistula (but otherwise physiologically normal) "the sight of and even the mere mention of food evoked contraction of the gall bladder" (p. 119). The same investigator also studied a patient with a fistula of the pancreatic duct but otherwise healthy and with a normal digestive tract; 1 or 2 minutes after being drawn into conversation about savory foods, this patient (who was kept on a special diet which served to inhibit secretions) "showed against this inhibitory background abundant pancreatic secretions." (The above patients had *not* participated in experimental conditioning procedures.)

Scantlebury and Patterson (1940) demonstrated that suggestions of eating a fictitious meal were effective in inducing a temporary and at times a complete cessation of gastric hunger contractions in a hypnotized *S*. Lewis and Sarbin (1943) repeated this experiment, employing the Carlson balloon-manometer method with eight *Ss* who had fasted prior to the

experiment. The *Ss* were first given the Friedlander-Sarbin hypnotic induction procedure and rated on "depth of hypnosis." Whenever the *Ss* showed gastric hunger contractions, they were given the suggestion of eating a meal. The kymographic tracings showed that the suggestions were effective in inhibiting the hunger contractions in the majority of trials with the "deeply hypnotized" *Ss*, in some of the trials with the "moderately hypnotized" *Ss*, and in none of the trials with *Ss* who were "slightly hypnotized" or not hypnotized. However, a comparable inhibition of hunger contractions could be demonstrated in the "deeply hypnotized" *Ss* by asking them to solve an arithmetic problem silently. No attempt was made to determine if hunger contractions could be inhibited in un hypnotized persons by asking them to "vividly imagine" eating a delicious meal.

Earlier studies which did not employ hypnotic procedures found comparable effects. For example, Carlson (1916, p. 152) found that after 4 days of fasting the sight and smell of food inhibited his hunger contractions. Since acid in contact with the gastric mucosa apparently acts reflexly to produce inhibition of gastric contractions (Carlson, 1916, pp. 175-176) and since the "mere thought" of appetizing food gives rise to a significant amount of hydrochloric acid secretion in some normal persons (Miller et al., 1920), it can be hypothesized that suggestions of eating a meal are effective in some "hypnotized" *Ss* and some un hypnotized *Ss* in inducing gastric acid secretions which act reflexly to inhibit the hunger movements.

In summary, the above studies on metabolic and gastrointestinal functions appear to indicate that blood

sugar levels can be altered in diabetic patients by hypnotic or nonhypnotic procedures which alter the level of "arousal," and gastric and pancreatic secretions and gastric hunger contractions can be influenced by symbolic stimulation in both hypnotized and un hypnotized persons.

EFFECT OF HYPNOTIC STIMULATION ON CUTANEOUS FUNCTIONS

Production of Herpetic Blisters (Cold Sores) by Hypnotic Stimulation

Ullman (1947) reported that a patient (who had been previously cured of hysterical blindness) showed multiple herpetic blisters on the lower lip 25 hours after it was suggested to him "while in hypnotic trance" that he appeared rundown and debilitated, he felt as if he were catching cold, and fever blisters were forming on his lower lip. Heilig and Hoff (1928) had previously demonstrated a similar effect in an experiment with three "neurotic" women. Their procedure was as follows: After a formal hypnotic induction, an intense emotional reaction was elicited from each *S* by suggesting an extremely unpleasant experience related to her previous life history. During the excitement, the experimenter stroked the *S*'s lower lip and suggested a feeling of itch such as she had experienced previously when a cold sore was forming. Within 48 hours after the termination of the experiment small blisters had appeared on the lower lip of each *S*. This report also includes the following data: at least two of the *S*s had a history of recurrent herpes labialis following emotional arousal; determination of the opsonic index before and after the hypnotic experiment indicated that the *S*'s physiological resistance was reduced after the experiment; herpetic blisters could not be induced

when the hypnotized *S*s were given direct suggestions that such blisters were forming without at the same time eliciting an emotional reaction.

The above studies can be placed in broader context by noting the following: (a) The herpes simplex virus appears to be ubiquitous and ready to produce illness whenever the normal balance between it and the host is disturbed not only by fever, allergic reactions, sunburn, and so forth, but also by emotional stress and by symbolic stimulation which has significance for the person (Sulzberger & Zardens, 1948). (b) Some persons show recurrent attacks of herpes simplex in the same localized area (Veress, 1936); in some cases the attacks appear to be closely related to "emotional conflicts" or to stimulation which tends to elevate the level of "arousal" (Blank & Brody, 1950; Schneck, 1947). These findings suggest that an "hypnotic induction" procedure and specific suggestions of blister formation may not be necessary to induce herpetic blisters in appropriately predisposed persons. An experiment along the following lines is indicated: An experimental group consisting of persons with a history of herpes labialis should be given appropriate stimulation to induce emotional arousal *without* an hypnotic procedure. A second experimental group consisting of persons who do not have a history of herpes should be placed in "deep hypnosis" and given specific suggestions of cold sore formation. It can be hypothesized that some of the un hypnotized *S*s in the first group will show herpetic blisters within a day or so after the experiment. It would be of interest to determine if any of the "deeply hypnotized" *S*s in the second group will show cold sores after the experiment.

Induction of Localized (Nonherpetic) Blisters by Hypnotic Stimulation

Pattie (1941) has reviewed 11 experiments which ostensibly demonstrate that localized blisters (excluding cold sores) can be evoked by direct suggestions given to somnambulistic hypnotic Ss. A relatively well controlled experiment reported by Hadfield (1917) can be taken as the prototype of these investigations: After the S was hypnotized, an assistant touched his arm while Hadfield gave continuous suggestions that a red-hot iron was being applied and that a blister would form in the burned area. The arm was then bound in a sealed bandage and the S was watched continuously during the following 24 hours. At the end of this period the bandage was opened in the presence of three physicians and, on the designated area, the beginning of a blister was noted which gradually developed during the day to form a large bleb surrounded by an area of inflammation. Although the other experiments followed this general pattern, there are numerous variations: in some instances, the experimenter stated that a blister would form after a definite time interval and in other instances no time was specified; some Ss were instructed to awaken immediately after the suggestion of bulla formation and others were not given such instructions until it was determined if the blister had formed; although in most instances the blister formed in the area specified, in at least two instances (Jendrassik, 1888; Smirnoff, 1912) the bleb formed in another body area. Also, in at least two experiments (Rybalkin, 1890; von Krafft-Ebing, 1889, pp. 26-27, 58-59) the controls were not satisfactory; the Ss were not observed during the intervening pe-

riod and it is possible that they may have deliberately injured the area.

Two additional cases have been reported since the publication of Pattie's (1941) review. Ullman's (1947) S, mentioned in the preceding section of the present paper, had previously been cured of hysterical blindness and had previously shown herpetic blisters after hypnotic stimulation. In an additional hypnotic session, the S was induced to recall the battle in which he had recently participated and was given the suggestion that a small particle of molten shell fragment had glanced off the dorsum of his hand. At this point in the procedure, the experimenter brushed the hand with a small flat file to add emphasis to the suggestion. Pallor followed immediately in this circumscribed area approximately 1 centimeter in diameter; after 20 minutes a narrow red margin was evident about the area of pallor and after 1 hour the beginning of a blister was noticeable. The S was then dismissed and returned approximately 4 hours later; at this time a bleb about 1 centimeter in diameter was evident. (The S was not observed during the intervening period.) More recently, Borelli and Geertz (Borelli, 1953) succeeded in inducing dermatological alterations which superficially resembled blister formation in a 27-year-old patient with "neurodermatitis." During "deep hypnosis" a coin was placed on the normal skin of the hand and it was suggested that a blister would form within a day at the spot where the fictitious burn was occurring. The next day the patient showed a sharply circumscribed and elevated area at the designated spot which superficially resembled a blister but could be more appropriately described as white dermographism.

With few if any exceptions investigators reporting positive results emphasize that they selected somnambulistic hypnotic Ss for their experiments; however, a number of workers using similar procedures with somnambulistic Ss have reported negative results in all cases (Sarbin, 1956; Wells, 1944), or have reported negative results with the majority of such Ss and positive results only in rare cases (Hadfield, 1920). These negative findings appear to indicate that appropriate suggestions given to "deeply hypnotized" persons may be necessary but by no means sufficient conditions for this phenomenon.

An additional factor which appears necessary is indicated by the following. The 13 persons who gave ostensibly positive results were not a cross section of the normal population: prior to the experiment, one had been cured of hysterical blindness and one had been cured of hysterical aphonia; during the time of the experiment, six were diagnosed as hysterical and one was being treated for "shell-shock." At least five of these Ss had histories of localized skin reactions: one had "a delicate skin" and showed labile vasomotor reactions (Doswald & Kreibich, 1906, Case 1), a second had suffered from "neurotic skin gangrene" and had a history of wheals following emotional arousal (Doswald & Kreibich, 1906, Case 2), a third had "a delicate skin" plus "dermographia of medium grade" (Heller & Schultz, 1909), a fourth had suffered from "hysterical ecchymoses" (Schindler, 1927), and a fifth was afflicted with atopic dermatitis (Borelli, 1953). This suggests that the induction of localized blisters by hypnotic stimulation may be possible only in a small group of persons with a unique physiological predisposition. What is the nature of this "predis-

position"? The data summarized below suggest a tentative answer.

Blister formation and wheal formation apparently involve similar physiological and biochemical processes: the circular wheal of urticaria, the linear wheals of dermographism, and the blister resulting from a burn can be viewed as variations of the "triple response" of the skin to injury, consisting of the release of histamine or a histamine-like substance such as 5-hydroxytryptamine (serotonin) from the Mast cells, a local dilation of the minute vessels, an increase in permeability of the vessels, and a widespread arteriolar dilation (Lewis, 1927; Nilzén, 1947). Nearly every type of stimulus that produces whealing when applied to the skin will lead to blistering if rendered more intense, and blister formation appears to differ from wheal formation primarily in that the increased permeability of the vessel walls is of greater degree, the transuded fluid typically forms a pool in the superficial layers of the skin, and the epidermal layers are gradually forced asunder (Lewis, 1927). This close relationship between wheals and blisters appears to be significant because of the following:

1. In at least two of the "successful" hypnotic experiments (Borelli, 1953; Doswald & Kreibich, 1906, Case 2) the dermatological changes induced were much more similar to wheals than to blisters.

2. A critical reading of the other reports suggests that the histological findings were rarely so clear-cut as to definitely conclude that blisters and not wheals were produced.

3. Some un hypnotized persons show localized wheals when recalling former experiences in which such dermatological effects occurred.

4. Some un hypnotized persons show localized wheals after mild me-

chanical stimulation.

Moody (1946, 1948) has presented two case studies of patients who developed localized wheals when recalling former experiences in which wheals occurred. The first patient had been previously hospitalized for sleepwalking with aggressive behavior. On one occasion, during this earlier hospitalization, the patient's hands had been tied behind his back during sleep and wheals had formed in the traumatized area. At a later time, when recalling this experience after hexobarbital administration, wheals appeared on both forearms in the area which had previously been tied. On at least 30 occasions when recalling earlier experiences of physical injury, the second patient (who was being treated for "nervous breakdown") showed swelling, bruising, and bleeding in the body parts where the original injury presumably occurred; for instance, when remembering a former occasion when she had been struck across the dorsum of both hands with a cutting whip, the patient showed wheals on both hands in the respective areas. Along similar lines, Graff and Wallerstein (1954) reported that during a therapeutic interview a 27-year-old sailor, who had a tattoo of a dagger on his arm, suddenly showed a wheal reaction sharply limited to the outline of the dagger. The wheal subsided after this session but reappeared again in the same way during a subsequent interview. The authors interpret the patient's free associations as indicating that the wheal had symbolic significance for the patient. Brandt (1950) has reported similar cases of patients showing sharply localized wheal reactions which appeared to be closely related to symbolic stimulation.

Dermographism (that is, wheal formation in response to a single mod-

erately strong stroking of the skin) is not as uncommon as is generally assumed. Testing 84 apparently normal young men, Lewis (1927) found a detectable swelling of the skin as a reaction to a single firm stroke in 25%; in 5% a full wheal developed. Some persons also show wheal formation at sites of mild pressure stimulation such as around a wristwatch strap, a belt, or a collar. Graham and Wolf (1950) reported an experimental study of 30 such persons who had a history of urticaria and showed "spontaneous" wheals in areas of mild pressure. All of these Ss also showed dermatographism although in some this was not apparent until stressful interviews had altered the condition of the skin vessels. Skin temperature measurements and indirect measurements of the state of the minute vessels (reactive hyperemia threshold) indicated that the Ss were prone to respond with vasodilation of both arterioles and minute vessels to numerous stimuli. Since in all but one of the successful hypnotic experiments tactual stimulation was employed to localize the pseudotrauma and since in many of the experiments the stimulus object was a small piece of metal and was either allowed to remain in contact with the skin or was replaced by a bandage, it appears plausible, as Weitzenhoffer (1953, p. 144) has suggested, that similar physiological mechanisms may be responsible for the above types of urticaria factitia and for at least some cases of the hypnotic production of localized "blisters."

The above data suggest an experiment as follows: Persons who show gross vasomotor alterations during seemingly slight changes in the stimulating situation or who show dermatographism under normal conditions or during stress should be given the

following instructions *without* a preliminary "hypnotic induction"—"Try to visualize a blister (in a specified area) and tell yourself repeatedly that such a blister is forming." If the Ss are adequately motivated to comply with these odd instructions, it can be hypothesized that some will show dermatological changes related to vesiculation. A second experimental group consisting of persons who do *not* show signs of vasomotor lability should be given suggestions of blister formation after an "hypnotic induction" procedure and when they appear to be in "the trance state." It would be of interest to determine if these "hypnotized" Ss will show any cutaneous reactions which are involved in the formation of a blister.

Cure of Warts by "Hypnosis"

Since the genesis of warts appears to be causally related to virus activity and since present day methods of treating warts are "roundabout and nonspecific" (Pillsbury, Shelley, & Kligman, 1956, p. 690), recent reports indicating that appropriate suggestions given to a hypnotized person are singularly effective in curing these benign epitheliomas are of unique interest. Asher (1956) found that suggestions of wart disappearance given to 25 hypnotizable patients resulted, after 4 to 20 treatments, in a complete cure in 15, a marked decrease in the number of warts in 4, and no apparent change in 6 patients. In these cases the warts before treatment varied from 2 to 53 and were present from 3 months to 6 years. Eight un hypnotizable patients given similar suggestions showed no diminution in the number of warts; however, in these cases the treatment was discontinued after 10 sessions. In a more extensive investigation, Ullman and Dodek (1960)

attempted to relieve warts by hypnotic suggestions in 62 adults attending an outpatient clinic. At weekly intervals each patient was given suggestions of sleep and drowsiness followed by suggestions to determine "the depth of hypnosis"; when the patient was judged to be at "the period of maximum hypnotic effect," he was told that the warts would begin to disappear. Of the 47 patients rated as "poor hypnotic subjects," only 2 showed wart regression within a 4-week period. However, 6 of the 15 patients rated as "good hypnotic subjects" had been cured of multiple common warts (or, in one case, of a single common wart) within 2 weeks following the initiation of treatment; within a 4-week period, 8 of the 15 showed wart involution. In these successful cases the mean duration of the warts prior to treatment was 19 months with a range of 3 weeks to 6 years.

The above investigations are open to the criticism that the warts may have shown spontaneous involution within the same period of time if no hypnotic treatment had been given. A recent study, however, appears to have satisfactorily controlled this factor. After an "hypnotic induction" consisting of eye fixation and suggestions of relaxation, Sinclair-Gieben and Chalmers (1959) suggested to 14 patients (with common warts present bilaterally for at least 6 months) that the warts on *one* side of the body would disappear. Ten of the 14 patients showed "adequate depth of hypnosis" as indicated by compliance with a simple posthypnotic suggestion and by partial or complete amnesia. Within 5 weeks to 3 months, 9 of these 10 "hypnotizable" patients showed wart involution on the "treated" side while the warts on the "control" side remained unchanged. (In one patient the "un-

treated" side showed wart regression 6 weeks after the "treated" side had been cured.) No benefit was observed from this treatment in the four patients who were not able to attain "adequate hypnotic depth."

Although the above studies indicate that symbolic stimulation is effective in inducing wart involution in some Ss who are able to attain "a deep hypnotic state," equally successful results have been reported for a variety of suggestive procedures which do not involve an "hypnotic induction" or "the trance state." Grumach (1927) found that 16 of 18 patients with longstanding warts showed complete regression of these structures within 1 to 4 months after being given, at intervals of 8 to 14 days, an intramuscular placebo injection (normal saline) in the upper arm while, at the same time, being told that they were receiving a new and powerful wart remedy. Allington (1934) followed-up 84 patients with longstanding warts treated with an intragluteal placebo injection (distilled water); 35 (or 41.7%) were relieved of plane warts or common warts after only one injection, 4 were cured after two injections, and 1 after three injections. Bloch (1927) reported comparable results with a somewhat different procedure. The patient was blindfolded and his hand was placed on a table containing an electric apparatus; although the electricity was started no current reached the patient. The warts were then painted with an innocuous dye, the blindfold was removed, and the patient, now confronted with the luridly colored warts, was told that the warts were dead and must not be washed until they had disappeared. Of 179 patients thus treated and adequately followed-up, 55 (or 30.7%) showed wart involution after the first session and an additional 43 patients (or

24%) showed wart involution after additional session extending over a period varying from 1 week to 3 months. Using similar procedures, Bonjour (1929), Sulzberger and Wolf (1934), and Vollmer (1946) reported success in a comparable percentage of cases with warts of from 2 to 6 years duration. In general, these suggestive procedures were more effective when the patient showed multiple warts rather than a single wart and when the warts were of the juvenile type rather than the common type; this type of treatment also tended to be more successful with recent lesions and with younger patients.

Would a similar percentage of patients have shown spontaneous remission of warts if they had not been "treated" in the specified period of time involved in the above experiments? Memmesheimer and Eisenlohr (1931) matched 70 patients treated by a suggestive procedure—painting the warts with methylene blue and suggesting their disappearance—with 70 patients with similar warts of similar duration not given any treatment. The results were as follows: at the end of 1 month, 11 of the treated patients showed wart resolution as compared to only 2 of the patients in the control group; at the end of 3 months, 14 of the treated patients were cured as compared to only 5 of the untreated; however, at the end of 6 months, 20 patients in the control group showed wart involution as compared to only 17 patients in the treated group. The conclusion suggested by this study, namely, that suggestive treatment may accelerate a spontaneous physiological process leading to wart involution, is supported by additional investigations summarized below.

Similar physiological processes have been demonstrated when warts

heal spontaneously and when they are cured in apparent response to symbolic stimulation. Unna (quoted by Samek, 1931) observed histologically that during spontaneous remission the normal cutis surrounding the wart showed a distinctive reaction consisting of hyperemia and cell proliferation. Other workers (Allington, 1952; Biberstein, 1944; Sulzberger & Wolf, 1934; Vollmer, 1946) have also noted a distinct inflammatory reaction immediately before spontaneous healing or before wart disappearance in apparent response to suggestion or to chemical treatment. In histological studies of warts undergoing involution in a patient treated by a suggestive procedure, Samek (1931) demonstrated a specific inflammatory reaction in the dermis consisting of dilation of blood vessels, hyperemia, edema, and perivascular infiltration of leucocytes (especially lymphocytes). Concomitant with this inflammatory reaction, mitoses became less frequent in the germinative epidermis (stratum mucosum); with mitoses almost at a standstill, the prickle-cell layer became thin, a normal stratum granulosum reformed, and the degenerated cells flaked off.

After a careful review of the above and related studies, Allington (1952) concluded that "at times the balance between susceptibility and immunity in warts must be a delicate one [and] only a slight shift may be needed to cause their disappearance." Vollmer (1946) had similarly concluded from an earlier review that a labile equilibrium must exist between the physiological processes which maintain the wart and those which cause wart involution and that appropriate verbal stimulation may alter the equilibrium in the direction of wart resolution by causing hyperemia in

the surrounding tissue. A number of earlier workers (Sulzberger & Wolf, 1934; Zwick, 1932) had also pointed to vasomotor changes as crucial factors in wart remission and, more recently, Ullman (1959) presented data indicating that when warts are treated by suggestion an "affective response" is induced in the patient and the mechanism of healing may be dependent on local vascular alterations which accompany the emotional reaction. Since a number of investigations reviewed in an earlier section of the present paper suggest that localized vasodilation and localized vasoconstriction can be induced in *some* individuals by symbolic stimulation—e.g., by asking the individual to recall former experiences in which such vasomotor alterations occurred (Menzies, 1941)—further investigations are required to determine the following: (a) Are local vasomotor changes consistently present when wart resolution is occurring after suggestive treatment? (b) If so, do such vasomotor effects accelerate a natural physiological process of wart remission? (c) Is treatment of warts by suggestive procedures relatively more successful in persons who show vasomotor lability, that is, in persons who respond with a greater than average degree of vasodilation or vasoconstriction to symbolic stimulation or to emotion-inducing stimulation?

THE PHYSIOLOGICAL CORRELATES OF "THE HYPNOTIC STATE"

The studies reviewed above suggest the general conclusion that many if not all of the physiological effects which can be induced in some Ss during "hypnosis" can also be induced in some persons without hypnosis. The experiments reviewed below suggest that it is difficult if not impossi-

ble to find a physiological index which differentiates "the hypnotic state" from "the normal waking state."

During recent years an extensive number of experiments have been designed to determine if "hypnosis" is characterized by an elevated or depressed metabolic rate, heart rate, blood pressure, skin conductance, respiratory rate, digital blood volume, etc. All of these investigations lead to a similar conclusion: Physiological functions vary in the same way during "hypnosis" as they do during "waking" behavior. Taking energy expenditure as the example, the evidence indicates that metabolic rate may be elevated, may be depressed, or may not be significantly altered during "the hypnotic state": Grafe and Mayer (1923) found that hypnotized Ss tended to show an elevated metabolic rate; von Eiff (1950) found that 16 Ss showed an average depression of 7% in "basal" metabolic rate during hypnosis; and Whitehorn, Lundholm, Fox, and Benedict (1932) reported that oxygen consumption was not significantly affected by hypnosis. Since the metabolic rate is elevated during "emotional arousal" and is depressed during relaxation and sleep (Best & Taylor, 1950, p. 622), these results are only superficially contradictory: Experimenters' finding that "hypnosis" depresses metabolism (von Eiff, 1950) had instructed their Ss to become relaxed, drowsy, and sleepy and had not given additional suggestions that could lead to arousal; investigators reporting that "the hypnotic state" does not affect metabolism (Whitehorn et al., 1932) had trained their Ss over a period of days to insure maximal relaxation when the metabolic rate was determined during the control experiment; experimenters

finding that heat production was elevated during "hypnosis" (Grafe & Mayer, 1923) had activated the Ss by suggesting various emotional experiences.

Investigations designed to determine if "hypnosis" is characterized by an elevated or depressed level of skin conductance have produced comparable results; during "the hypnotic trance" Ss may show an elevation, a slight decrease, or no significant change in palmar conductance (Barber & Coules, 1959; Davis & Kantor, 1935; Estabrooks, 1930; Levine, 1930). Since an elevated conductance level generally indicates an elevated "activation" level and a low level of conductance generally indicates a low level of "arousal" (Duffy, 1957; Woodworth & Schlosberg, 1954), these results are in agreement as follows: (a) Hypnotic Ss show an elevated level of palmar conductance when they *carry out* suggestions which involve effort or activity (Barber & Coules, 1959; Davis & Kantor, 1935). (b) When given suggestions of relaxation and drowsiness, Ss participating in hypnotic experiments may show a decrease or no significant change in palmar conductance; if the S accepts the suggestions literally and relaxes, he shows a fall in conductance (Davis & Kantor, 1935; Estabrooks, 1930; Levine, 1930); if the S is aware that suggestions of drowsiness and relaxation are not meant to be taken literally, i.e., if he has learned from previous participation in hypnotic experiments that to carry out subsequent suggestions properly he must remain alert, he generally shows no significant change in conductance (Barber & Coules, 1959). Investigations along similar lines which support the general conclusion that the "hypnotized" person does not differ signifi-

cantly from the normal person in heart rate, respiratory rate, blood pressure, digital blood volume, etc. have been reviewed by Gorton (1949) Weitzenhoffer (1953), Sarbin (1956), and Crasilneck and Hall (1959).

Some years ago it seemed that the electroencephalograph would prove to be a valuable tool for determining when a person was or was not "hypnotized." This hope has not been realized. Extensive work in this area, reviewed by Weitzenhoffer (1953) and Chertok and Kramarz (1959), has demonstrated that in the great majority of instances the hypnotized person continues to show his characteristic waking pattern on the EEG. However, if the operator makes it clear to the *S* that he should actually sleep—for example, by not disturbing the *S* after instructing him to sleep—some *Ss* participating in hypnotic experiments show delta activity on the EEG, indicating that they have literally gone to sleep (Barker & Burgwin, 1948; Schwarz et al., 1955), and others show "periods of brief flattening out of the record . . . sometimes accompanied by infrequent isolated theta rhythms," indicating that they have gone into a light sleep (Chertok & Kramarz, 1959, p. 233). However, when the *S* is once more stimulated verbally by the hypnotist, he again shows his characteristic waking pattern on the EEG. In brief, studies employing the EEG indicate that the "hypnotized" person remains normally awake until it is made clear that he should literally go to sleep and is then permitted to sleep.

Within recent years, Lovett Doust (1953) and Ravitz (1951, 1959) have proposed two additional physiological indices of "the hypnotic state." Employing three hysterics and one psychopath as *Ss*, Lovett Doust found that "the induction of hypnosis" was

consistently accompanied by a significant fall in arterial oxygen saturation levels as measured by discontinuous spectroscopic oximetry at the fingernail fold. However, the term "induction of hypnosis," as used in this report, does *not* imply that the *Ss* carried out one or more of the classical hypnotic behaviors, e.g., limb rigidities, negative or positive hallucinations; on the contrary, by this term the author refers to no more than the following: After being given suggestions of drowsiness, lethargy, and sleep, the *Ss* appeared passive and lethargic. Since a person who appears passive and lethargic is not necessarily "hypnotized" (that is, does not necessarily carry out any of the classical hypnotic behaviors) and since a person who carries out all of the classical hypnotic behaviors does not necessarily appear drowsy or passive (Barber, 1960b; Barber & Coules, 1959; Wells, 1924), Lovett Doust's findings are open to the following interpretation: A relative anoxemia is found during drowsiness or passivity and is not necessarily found when a person is "in the hypnotic state," i.e., when he carries out the classical hypnotic behaviors. Supporting evidence for this interpretation is presented in a previous study by the same investigator (Lovett Doust & Schneider, 1952) which demonstrated a similar fall in oximetric values during sleep.

Measuring standing potentials between the forehead and the palm of the hand, Ravitz (1951, 1959) found that the "hypnotic induction" procedure was accompanied by either a gradual increase or decrease in mean potential and "the trance state itself, following induction" was typically characterized by a voltage decrease and by an increased regularity of the direct current (DC) tracings. However, additional data presented in the

reports suggest that a decrease in voltage and an increased regularity of the DC tracings may be present whenever a person is relaxed and shows a low "arousal" level; for example, Ravitz notes that a decrease in voltage and an increase in regularity of the tracings are found during sleep and that increased voltage and decreased regularity are found during "changes in energy level," excitability, loquaciousness, grief, anxiety, and so forth. Since, as pointed out above and as will be discussed further below, a *S* need not show relaxation or passivity when he carries out the classical hypnotic behaviors, no conclusions can be deduced from these findings until the following hypotheses are experimentally confirmed or disproved: (a) Unhypnotized *Ss*, i.e., *Ss* who do *not* carry out such behaviors as limb rigidity, negative and positive hallucinations, age-regression, or amnesia when given appropriate suggestions, show a relative decrease in voltage and an increased regularity of the DC tracings when instructed to become relaxed and passive. (b) If an "hypnotic induction" leading to drowsiness and passivity is not employed, if, on the contrary, a direct suggestive procedure is used as described by Wells (1924) and Barber (1960b), "deeply hypnotized" *Ss*, i.e., *Ss* who carry out all of the classical hypnotic behaviors, do *not* show the above indices of "the trance state."

The above investigations and more recent speculations concerning the neurophysiological correlates of "hypnosis" (Arnold, 1959; Roberts, 1960; West, 1960) appear to be based on the following implicit assumptions: (a) When a person carries out the type of behavior which has been historically associated with the term "hypnosis" he is in "an altered state" from his normal self, specifically, in

"a trance state" or "an hypnotic state." (b) This "altered state" is of such a kind as to include a distinct and consistent type of physiological functioning which is rarely if ever present when a person is not carrying out hypnotic behavior. Although these assumptions are by no means limited to recent investigations (they are present in many if not all theories of "hypnosis" since the days of Mesmer), the evidence summarized below suggests that they are open to question.

HYPNOTIC BEHAVIOR WITHOUT AN "HYPNOTIC INDUCTION"

Since *Ss* participating in hypnotic experiments are almost always given an "hypnotic induction" consisting of suggestions of relaxation, drowsiness, and sleep, and since such a procedure is generally effective in inducing an appearance of lethargy or "trance," it often seems as if hypnotic behavior is a function of, or closely related to, "the trance state." However, in a pioneering study, Wells (1924) demonstrated that direct commands (e.g., "Your arm is insensitive to pain," "You cannot speak your name"), repeated emphatically for a few seconds, were sufficient to elicit anesthesia, amnesia for name, limb rigidity, hallucinatory pain, total amnesia, automatic writing, and posthypnotic behavior in a large proportion of male college students. Wells insisted that his *Ss* did not appear relaxed, drowsy, or lethargic and that he obtained results more quickly and with a larger proportion of *Ss* by such a direct procedure than by an "hypnotic induction" designed to induce "trance."

Recent investigations appear to confirm Wells' results. In one study (Barber, 1960b) a female student research assistant (untrained as a hypnotist) gave 236 students at a

girl's college direct suggestions (each suggestion requiring either 30 or 45 seconds) of body immobility—"Your body is heavy, rigid, solid; it's impossible for you to stand up; try, you can't,"—arm heaviness, arm levitation, hand rigidity, inability to say name, hallucination of thirst, selective amnesia, and posthypnotic behavior. Although an "hypnotic induction" was not employed, 49 Ss (or 20.8%) immediately carried out at least six of the eight suggestions and a total of 109 Ss (or 46%) carried out at least half of the suggestions. The postexperimental reports of these Ss were indistinguishable from the reports of persons who are said to be "hypnotized," e.g., "I just couldn't get up from the chair," "I was amazed when I couldn't speak my name," "I felt I was dying from thirst." In another study (Barber, 1960b, 1960c) the results of such a direct procedure were compared with the results of a formal "hypnotic induction" procedure. In the first part of this experiment 70 attendants, nurses, and clerical workers at a state hospital (who agreed to participate in an experiment on "imagination") were given a series of suggestions (each suggestion requiring 30 seconds) appropriate to induce arm rigidity, arm levitation, limb heaviness, limb anesthesia, hallucinations of thirst, heat, and cold, eye catalepsy, and hypnotic dream. Similar results were obtained as in the above study: 20 Ss (or 28.6%) immediately carried out at least seven of the nine suggestions and a total of 34 Ss (or 48.6%) carried out at least five of the suggestions. In the second part of this experiment the same Ss were given an "hypnotic induction" procedure (consisting of suggestions of relaxation, drowsiness, and sleep) and then given the suggestions of arm rigidity, arm levitation, limb anes-

thesia, etc., as in the preceding experiment. Although the Ss now appeared to be "in trance" (and stated, after the experiment, that they had "felt hypnotized") a high correlation ($r=.84$) was obtained between scores in the two sessions; in general, Ss who carried out one or two suggestions in the first part of the experiment carried out the same one or two suggestions after the "hypnotic induction" and Ss who responded positively to all of the suggestions in the second part of the experiment had also carried out all of the suggestions without the "hypnotic induction."

Related to the above are the results of other recent investigations (Barber 1958a; Fisher, 1954) which indicate the following:

1. If Ss participating in "hypnosis" experiments show lethargy, drowsiness, or other signs of "trance," these characteristics can be readily removed and the "good" Ss will continue to carry out the hypnotic performances if instructed: "Be perfectly awake. Come out of 'trance' but continue to obey my commands."

2. Many if not all "good" hypnotic Ss carry out all suggestions given during the posthypnotic period, i.e., after they are told to wake up, as long as they believe that their relationship with the operator remains that of subject and hypnotist.

In brief: Investigations which propose to find the physiological correlates of "hypnosis" uncritically assume that hypnotic behavior is a function of "the trance state"; this assumption is open to question. Appropriately predisposed persons do not need an "hypnotic induction" and need not appear to be in "trance" to carry out many if not all of the behaviors which have been associated with the term "hypnosis."

SUMMARY AND CONCLUSIONS

1. The normal person who is asked to "concentrate away from" red and green responds to the Ishihara in the manner characteristic of the hypnotic "color-blind" subject. An "hypnotic induction" procedure and "the trance state" may also be superfluous to eliciting the behavior which characterizes "hypnotic blind" or "hypnotic deaf" subjects; the evidence reviewed suggests that similar performances can be induced in normal persons by simply instructing them to remain inattentive and unresponsive to visual or auditory stimuli.

2. A number of physiological effects which have been considered as peculiar to "the hypnotic state" appear to be relatively commonplace performances; e.g., although suggestions of eating a delicious meal are at times effective in evoking gastric and pancreatic secretions and in inhibiting gastric hunger contractions in some "deeply hypnotized" subjects, it is not uncommon for normal persons to show similar gastrointestinal effects when they visualize the ingestion of savory food.

3. A group of so-called "hypnotic" phenomena—production of localized blisters, cure of warts, alteration of blood glucose levels, production of

tachycardia or cardiac block—can apparently be elicited with or without an "hypnotic induction" in a small number of individuals who possess a specific lability of the physiological systems involved.

4. An extensive series of experiments has failed to find a physiological index which differentiates "the hypnotic state" from "the waking state."

5. A series of experiments comparing the results of an "hypnotic induction" procedure with the results of a direct suggestive procedure indicate that appropriately predisposed persons do not need an "hypnotic induction" and need not appear to be in "the trance state" to carry out the typical behaviors which have been associated with the word "hypnosis."

6. Further investigations into the nature of "hypnosis" might well bypass the concepts of "hypnotic induction" and "trance state" and focus on biographical and situational factors which may account for certain individuals responding to symbolic stimulation from another person with so-called "hypnotic" behavior, whether primarily motor responses (e.g., limb rigidity, eye catalepsy) or primarily physiological responses (e.g., tachycardia, wart involution).

REFERENCES

- ALLINGTON, H. V. Sulpharsphenamine in the treatment of warts. *Arch. Dermatol. Syphilol.*, NY, 1934, **29**, 687-690.
- ALLINGTON, H. V. Review of psychotherapy of warts. *Arch. Dermatol. Syphilol.*, NY, 1952, **66**, 316-326.
- ARNOLD, MAGDA B. Brain function in hypnosis. *Int. J. clin. exp. Hypnosis*, 1959, **7**, 109-119.
- ASHER, R. Respectable hypnosis. *Brit. med. J.*, 1956, **1**, 309-313.
- BARBER, T. X. Hypnosis as perceptual-cognitive restructuring: II. "Post"-hypnotic behavior. *J. clin. exp. Hypnosis*, 1958, **6**, 10-20. (a)
- BARBER, T. X. Hypnosis as perceptual-cognitive restructuring: IV. "Negative hallucinations." *J. Psychol.*, 1958, **46**, 187-201. (b)
- BARBER, T. X. Toward a theory of pain: Relief of chronic pain by prefrontal leucotomy, opiates, placebos, and hypnosis. *Psychol. Bull.*, 1959, **56**, 430-460.
- BARBER, T. X. "Hypnosis," analgesia, and the placebo effect. *J. Amer. Med. Ass.*, 1960, **172**, 680-683. (a)
- BARBER, T. X. The necessary and sufficient conditions for hypnotic behavior. *Amer. J. clin. Hypnosis*, 1960, **3**, 31-42. (b)
- BARBER, T. X. The reality of hypnotic be-

- havior and the fiction of "hypnosis." Paper read at American Psychological Association, Chicago, September 1960. (c)
- BARBER, T. X., & COULES, J. Electrical skin conductance and galvanic skin response during "hypnosis." *Int. J. clin. exp. Hypnosis*, 1959, 7, 79-92.
- BARBER, T. X., & DEELEY, D. C. Experimental evidence for a theory of hypnotic behavior: I. "Hypnotic color-blindness" without "hypnosis." *Int. J. clin. exp. Hypnosis*, 1961, 9, 79-86.
- BARKER, W., & BURGWIN, SUSAN. Brain wave patterns accompanying changes in sleep and wakefulness during hypnosis. *Psychosom. Med.*, 1948, 10, 317-326.
- BAUCH, M. Beeinflussung des Diabetes mellitus durch psychophysische Entspannungsübungen. *Dtsch. Arch. klin. Med.*, 1935, 2, 149-166.
- BENNETT, L. L., & SCOTT, N. E. The production of electrocardiographic abnormalities by suggestion under hypnosis: A case report. *Amer. Practit.*, 1949, 4, 189-190.
- BERMAN, R., SIMONSON, E., & HERON, W. Electrocardiographic effects associated with hypnotic suggestion in normal and coronary sclerotic individuals. *J. appl. Physiol.*, 1954, 7, 89-92.
- BEST, C. H., & TAYLOR, N. B. *The physiological basis of medical practice*. (5th ed.) Baltimore: Williams & Wilkins. 1950.
- BIBERSTEIN, H. Immunization therapy of warts. *Arch. Dermatol. Syphilol.*, NY, 1944, 50, 12-22.
- BLANK, H., & BRODY, M. Recurrent herpes simplex: A psychiatric and laboratory study. *Psychosom. Med.*, 1950, 12, 254-260.
- BLOCH, B. Ueber die Heilung der Warzen durch Suggestion. *Klin. Wschr.*, 1927, 6, 2271-2275, 2320-2325.
- BONJOUR, J. Influence of the mind on the skin. *Brit. J. Dermatol.*, 1929, 41, 324-326.
- BORELLI, S. Psychische Einflüsse und reaktive Hauterscheinungen. *Munch. med. Wschr.*, 1953, 95, 1078-1082.
- BRANDT, R. A tentative classification of psychological factors in the etiology of skin diseases. *J. invest. Dermatol.*, 1950, 14, 81-90.
- BYKOV, K. M. *The cerebral cortex and the internal organs*. New York: Chemical Publishing, 1957.
- CALLAHAN, A., & REDLICH, F. C. Electroencephalography and ophthalmology. *Amer. J. Ophthal.*, 1946, 29, 1522-1533.
- CARLSON, A. J. *The control of hunger in health and disease*. Chicago: Univ. Chicago Press, 1916.
- CHAPMAN, L. F., GOODELL, HELEN, & WOLFF, H. G. Increased inflammatory reaction induced by central nervous system activity. *Trans. Ass. Amer. Physicians*, 1959, 72, 84-109.
- CHERTOK, L., & KRAMARZ, P. Hyponosis, sleep and electro-encephalography. *J. nerv. ment. Dis.*, 1959, 128, 227-238.
- CRASILNECK, H. B., & HALL, J. A. Physiological changes associated with hypnosis: A review of the literature since 1948. *Int. J. clin. exp. Hypnosis*, 1959, 7, 9-50.
- CREDE, R. H., CHIVERS, N. C., & SHAPIRO, A. P. Electrocardiographic abnormalities associated with emotional disturbances. *Psychosom. Med.*, 1951, 13, 277-288.
- DAVIS, R. C., & KANTOR, J. R. Skin resistance during hypnotic states. *J. gen. Psychol.*, 1935, 13, 62-81.
- DELHOUGNE, R., & HANSEN, K. Die suggestive Beeinflussbarkeit der Magen- und Pankreassekretion in der Hypnose. *Dtsch. Archiv. klin. Med.*, 1927, 157, 20-35.
- DORCUS, R. M. Modification by suggestion of some vestibular and visual responses. *Amer. J. Psychol.*, 1937, 49, 82-87.
- DOSWALD, D. C., & KREIBICH, K. Zur Frage der posthypnotischen Hautphänomene. *Mh. prakt. Dermatol.*, 1906, 43, 634-640.
- DUFFY, ELIZABETH. The psychological significance of the concept of "arousal" or "activation." *Psychol. Rev.*, 1957, 64, 265-275.
- DUNBAR, FLANDERS. *Emotions and bodily changes*. (4th ed.) New York: Columbia Univ. Press, 1954.
- DYNES, J. B. An experimental study in hypnotic anesthesia. *J. abnorm. soc. Psychol.*, 1932, 27, 79-88.
- ERICKSON, M. H. A study of clinical and experimental findings on hypnotic deafness: I. Clinical experimentation and findings. *J. gen. Psychol.*, 1938, 19, 127-150. (a)
- ERICKSON, M. H. A study of clinical and experimental findings on hypnotic deafness: II. Experimental findings with a conditioned response technique. *J. gen. Psychol.*, 1938, 19, 151-167. (b)
- ERICKSON, M. H. The induction of color blindness by a technique of hypnotic suggestion. *J. gen. Psychol.*, 1939, 20, 61-89.
- ESTABROOKS, G. H. The psychogalvanic reflex in hypnosis. *J. gen. Psychol.*, 1930, 3, 150-157.
- FAVILL, J., & WHITE, P. D. Voluntary acceleration of the rate of the heart beat. *Heart*, 1917, 6, 175-188.
- FISHER, S. The role of expectancy in the performance of posthypnotic behavior. *J. abnorm. soc. Psychol.*, 1954, 49, 503-507.
- FISHER, V. E. Hypnotic suggestion and the

- conditioned reflex. *J. exp. Psychol.*, 1932, 15, 212-217.
- FORD, W. L., & YEAGER, C. L. Changes in electroencephalogram in subjects under hypnosis. *Dis. nerv. Syst.*, 1948, 9, 190-192.
- GERARD, R. W. General discussion of symposium. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior: The Hixon symposium*. New York: Wiley, 1951. P. 94.
- GIGON, A., AIGNER, E., & BRAUCH, W. Ueber den Einfluss der Psyche auf körperliche Vorgänge: Hypnose und Blutzucker. *Schweiz. med. Wschr.*, 1926, 56, 749-750.
- GORTON, B. E. The physiology of hypnosis. *Psychiat. Quart.*, 1949, 23, 317-343, 457-485.
- GRAFE, E., & MAYER, L. Ueber den Einfluss der Affekte auf den Gesamtstoffwechsel. *Z. ges. Neurol. Psychiat.*, 1923, 86, 247-253.
- GRAFF, N. I., & WALLERSTEIN, R. S. Unusual wheal reaction in a tattoo. *Psychosom. Med.*, 1954, 16, 505-515.
- GRAHAM, D. T., & WOLF, S. Pathogenesis of urticaria: Experimental study of life situations, emotions, and cutaneous vascular reactions. *J. Amer. Med. Ass.*, 1950, 143, 1396-1402.
- GREYER, W. F. A comment on "The induction of color blindness by a technique of hypnotic suggestion." *J. gen. Psychol.*, 1940, 23, 207-210.
- GRUMACH, L. Ueber Suggestivbehandlung von Warzen. *Munch. med. Wschr.*, 1927, 74, 1093-1094.
- HADFIELD, J. A. The influence of hypnotic suggestion on inflammatory conditions. *Lancet*, 1917, 2, 678-679.
- HADFIELD, J. A. The influence of suggestion on body temperature. *Lancet*, 1920, 2, 68-69.
- HARRIMAN, P. L. Hypnotic induction of color vision anomalies: I. The use of the Ishihara and the Jensen tests to verify the acceptance of suggested color blindness. *J. gen. Psychol.*, 1942, 27, 289-298.
- HEILIG, R., & HOFF, H. Ueber Psychogene Entstehung des Herpes labialis. *Med. Klin.*, 1928, 24, 1472.
- HELLER, F., & SCHULTZ, J. H. Ueber einen Fall hypnotisch erzeugter Blasenbildung. *Munch. med. Wschr.*, 1909, 56, 2112.
- HERNÁNDEZ-PEÓN, R., & DONOSO, M. Influence of attention and suggestion upon subcortical evoked electrical activity in the human brain. In L. van Bogaert & J. Radermecker (Eds.), *First International Congress of Neurological Sciences*. Vol. III. London: Pergamon, 1959. Pp. 385-396.
- HEYER, G. R. Psychogene Funktionsstörungen des Verdauungstraktes. In O. Schwarz (Ed.), *Psychogenese und Psychotherapie körperlicher Symptome*. Wien: Springer, 1925. Pp. 229-257.
- HILGARD, E. R., & MARQUIS, D. G. *Conditioning and learning*. New York: Appleton-Century, 1940.
- HINKLE, L. E., JR., & WOLF, S. A summary of experimental evidence relating life stress to diabetes mellitus. *J. Mt. Sinai Hosp.*, 1953, 19, 537-570.
- HUTTENLOCHER, J., & WESTCOTT, M. R. Some empirical relationships between respiratory activity and heart rate. *Amer. Psychologist*, 1957, 12, 414. (Abstract)
- IKEMI, Y., AKAGI, M., MAEDA, J., FUKUMOTO, T., KAWATE, K., HIRAKAWA, K., CONDO, S., NAKAGAWA, T., HONDA, T., SAKAMOTO, A., & KUMAGAI, M. Hypnotic experiments on the psychosomatic aspects of gastrointestinal disorders. *Int. J. clin. exp. Hypnosis*, 1959, 7, 139-150.
- JASPER, H. H., & CARMICHAEL, L. Electrical potentials from the intact human brain. *Science*, 1935, 81, 51-53.
- JASPER, H. H., & CRUIKSHANK, RUTH M. Electroencephalography: II. Visual stimulation and the after-image as affecting the occipital alpha rhythm. *J. gen. Psychol.*, 1937, 17, 29-48.
- JENDRASSIK, E. Einiges über suggestion. *Neurol. Zbl.*, 1888, 7, 281-283, 321-330.
- JENNESS, A., & WIBLE, C. L. Respiration and heart action in sleep and hypnosis. *J. gen. Psychol.*, 1937, 16, 197-222.
- KING, J. T., JR. An instance of voluntary acceleration of the pulse. *Bull. Johns Hopkins Hosp.*, 1920, 31, 303-304.
- KLINE, M. V., GUZE, H., & HAGGERTY, A. D. An experimental study of the nature of hypnotic deafness: Effects of delayed speech feedback. *J. clin. exp. Hypnosis*, 1954, 2, 145-156.
- KOEHLE, M. Ueber die willkürliche Beschleunigung des Herzschlages beim Menschen. *Arch. ges. Physiol.*, 1914, 158, 579-622.
- LANDIS, C., & SLIGHT, D. Studies of emotional reactions: VI. Cardiac responses. *J. gen. Psychol.*, 1929, 2, 413-420.
- LEVINE, M. Electrical skin resistance during hypnosis. *Arch. Neurol. Psychiat.*, 1930, 24, 937-942.
- LEWIS, J. H., & SARBIN, T. R. Studies in psychosomatics: I. The influence of hypnotic stimulation on gastric hunger contractions. *Psychosom. Med.*, 1943, 5, 125-131.
- LEWIS, T. *The blood vessels of the human skin and their responses*. London: Shaw, 1927.
- LJUNG, O. Ekg-Koronarinsuffizienz bei vegetativer Labilität. *Cardiologia*, 1949, 14, 191-218.

- LOFTUS, T. A., GOLD, H., & DIETHELM, O. Cardiac changes in the presence of intense emotion. *Amer. J. Psychiat.*, 1945, **101**, 697-698.
- LOOMIS, A. L., HARVEY, E. N., & HOBART, G. Electrical potentials of the human brain. *J. exp. Psychol.*, 1936, **19**, 249-279.
- LOVETT DOUST, J. W. Studies on the physiology of awareness: Oximetric analysis of emotion and the differential planes of consciousness seen in hypnosis. *J. clin. exp. Psychopathol.*, 1953, **14**, 113-126.
- LOVETT DOUST, J. W., & SCHNEIDER, R. A. Studies on the physiology of awareness: Anoxia and the levels of sleep. *Brit. med. J.*, 1952, **1**, 449-453.
- LUCKHARDT, A. B., & JOHNSTON, R. L. Studies in gastric secretions: I. The psychic secretion of gastric juice under hypnosis. *Amer. J. Physiol.*, 1924, **70**, 174-182.
- LUNDHOLM, H. An experimental study of functional anesthetics as induced by suggestions in hypnosis. *J. abnorm. soc. Psychol.*, 1928, **23**, 337-355.
- LUNDHOLM, H., & LOWENBACH, H. Hypnosis and the alpha activity of the electroencephalogram. *Charac. Pers.*, 1942, **11**, 144-149.
- MCCLURE, C. M. Cardiac arrest through volition. *Calif. Med.*, 1959, **90**, 440-441.
- MCDOWELL, M. Hypnosis in dermatology. In J. M. Schneck (Ed.), *Hypnosis in modern medicine*. (2nd ed.) Springfield, Ill.: Charles C Thomas, 1959. Pp. 101-115.
- MADOW, L. Cortical blindness. *J. Neuropathol.*, 1958, **17**, 324-332.
- MAINZER, F., & KRAUSE, M. The influence of fear on the electrocardiogram. *Brit. heart J.*, 1940, **2**, 221-230.
- MALMO, R. B., BOAG, T. J., & RAGINSKY, B. B. Electromyographic study of hypnotic deafness. *J. clin. exp. Hypnosis*, 1954, **2**, 305-317.
- MALMO, R. B., DAVIS, J. F., & BARZA, S. Total hysterical deafness: An experimental case study. *J. Pers.*, 1952, **21**, 188-204.
- MARCUS, H., & SAHLGREN, E. Untersuchungen über die Einwirkung der Hypnotischen Suggestion auf die Funktion des vegetativen Systems. *Münch. med. Wschr.*, 1925, **72**, 381-382.
- MEMMESHEIMER, A. M., & EISENLOHR, E. Untersuchungen über die Suggestivebehandlung der Warzen. *Dermatol. Z.*, 1931, **62**, 63-68.
- MENZIES, R. Further studies of conditioned vasomotor responses in human subjects. *J. exp. Psychol.*, 1941, **29**, 457-482.
- MILLER, R. J., BERGHEIM, O., REHFUSS, M. E., & HAWK, P. B. Gastric response to food: X. The psychic secretion of gastric juice in normal men. *Amer. J. Physiol.*, 1920, **52**, 1-27.
- MIRSKY, I. A. Emotional factors in the patient with diabetes mellitus. *Bull. Menninger Clin.*, 1948, **12**, 187-194.
- MOHR, F. *Psychophysische Behandlungsmethoden*. Leipzig: Hirzel, 1925.
- MOODY, R. L. Bodily changes during abreaction. *Lancet*, 1946, **2**, 934-935.
- MOODY, R. L. Bodily changes during abreaction. *Lancet*, 1948, **1**, 964.
- NIELSEN, O. J., & GEERT-JORGENSEN, E. Untersuchungen über die Einwirkung der hypnotischen Suggestion auf den Blutzucker bei Nichtdiabetikern. *Klin. Wschr.*, 1928, **7**, 1467-1468.
- NILZÉN, A. Studies in histamine. *Acta dermat.-venereol.*, Stockholm, 1947, **27**, Suppl. 17, 1-67.
- OGDEN, E., & SHOCK, N. W. Voluntary hypercirculation. *Amer. J. med. Sci.*, 1939, **198**, 329-342.
- PATTIE, F. A. A report of attempts to produce unocular blindness by hypnotic suggestion. *Brit. J. med. Psychol.*, 1935, **15**, 230-241.
- PATTIE, F. A. The production of blisters by hypnotic suggestions: A review. *J. abnorm. soc. Psychol.*, 1941, **36**, 62-72.
- PATTIE, F. A. The genuineness of unilateral deafness produced by hypnosis. *Amer. J. Psychol.*, 1950, **63**, 84-86.
- PEASE, E. A. Voluntary control of the heart. *Boston med. surg. J.*, 1889, **120**, 525-529.
- PILLSBURY, D. M., SHELLEY, W. B., & KLIGMAN, A. M. *Dermatology*. Philadelphia: Saunders, 1956.
- POVORINSKI, J. A., & FINNE, W. N. Der Wechsel des Zuckergehalts des Blutes unter dem Einfluss einer hypnotisch suggerierten Vorstellung. *Z. ges. Neurol. Psychiat.*, 1930, **129**, 135-146.
- RAGINSKY, B. B. Temporary cardiac arrest induced under hypnosis. *Int. J. clin. exp. Hypnosis*, 1959, **7**, 53-68.
- RAVITZ, L. J. Standing potential correlates of hypnosis and narcosis. *Arch. Neurol. Psychiat.*, 1951, **65**, 413-436.
- RAVITZ, L. J. Application of the electrodynamic field theory in biology, psychiatry, medicine, and hypnosis. *Amer. J. clin. Hypnosis*, 1959, **1**, 135-150.
- ROBERTS, D. R. An electrophysiological theory of hypnosis. *Int. J. clin. exp. Hypnosis*, 1960, **8**, 43-55.
- RYBALKIN, J. Brûlure du second degré provoquée par suggestion. *Rev. Hypnot.*, Paris, 1890, **4**, 361-362.
- SAMEK, J. Zum wesen der Suggestiven War-

- zenheilung. *Dermatol. Wschr.*, 1931, 93, 1853-1857.
- SARBIN, T. R. Physiological effects of hypnotic stimulation. In R. M. Dorcus (Ed.), *Hypnosis and its therapeutic applications*. New York: McGraw-Hill, 1956. Ch. 4.
- SCANTLEBURY, R. E., & PATTERSON, T. L. Hunger motility in a hypnotized subject. *Quart. J. exp. Physiol.*, 1940, 30, 347-358.
- SCHINDLER, R. *Nervensystem und spontane Blutunge*. Berlin: Karger, 1927.
- SCHNECK, J. M. Psychogenic component in a case of herpes simplex. *Psychosom. Med.*, 1947, 9, 62-64.
- SCHULTZ, J. H., & LUTHE, W. *Autogenic training*. New York: Grune & Stratton, 1959.
- SCHWARZ, B. E., BICKFORD, R. G., & RASMUSSEN, W. C. Hypnotic phenomena, including hypnotically activated seizures, studied with the electroencephalogram. *J. nerv. ment. Dis.*, 1955, 122, 564-574.
- SINCLAIR-GIEBEN, A. H. C., & CHALMERS, D. Evaluation of treatment of warts by hypnosis. *Lancet*, 1959, 2, 480-482.
- SMIRNOFF, D. Zur Frage der durch hypnotische Suggestion hervorgerufenen vasomotorischen Störungen. *Z. Psychother. med. Psychol.*, 1912, 4, 171-175.
- SOLOVEY, GALINA, & MILECHNIN, A. Concerning the nature of hypnotic phenomena. *J. clin. exp. Hypnosis*, 1957, 5, 67-76.
- SULZBERGER, M. B., & WOLF, J. The treatment of warts by suggestion. *Med. Rec.*, NY, 1934, 140, 552-557.
- SULZBERGER, M. B., & ZARDENS, S. H. Psychogenic factors in dermatologic disorders. *Med. Clin. N. Amer.*, 1948 (May), 669-685.
- TARCHANOFF, J. R. Ueber die willkürliche Acceleration der Herzschläge beim Menschen. *Arch. ges. Physiol.*, 1885, 35, 109-137.
- TAYLOR, N. B., & CAMERON, H. G. Voluntary acceleration of the heart. *Amer. J. Physiol.*, 1922, 61, 385-398.
- ULLMAN, M. Herpes simplex and second degree burn induced under hypnosis. *Amer. J. Psychiat.*, 1947, 103, 828-830.
- ULLMAN, M. On the psyche and warts: I. Suggestions and warts: A review and comment. *Psychosom. Med.*, 1959, 21, 473-488.
- ULLMAN, M., & DODEK, STEPHANIE. On the psyche and warts: II. Hypnotic suggestion and warts. *Psychosom. Med.*, 1960, 22, 68-76.
- VAN DE VELDE, T. H. Ueber willkürliche Vermehrung der Pulsfrequenz beim Menschen. *Arch. ges. Physiol.*, 1897, 66, 232-240.
- VAN PELT, S. J. The control of heart rate by hypnotic suggestion. In L. M. LeCron (Ed.), *Experimental hypnosis*. New York: Macmillan, 1954. Pp. 268-275.
- VERESS, F. V. Beiträge zur Pathogenese des Herpes simplex. *9th Int. Conv. Dermatologists*, 1936, 2, 242.
- VOLLMER, H. Treatment of warts by suggestion. *Psychosom. Med.*, 1946, 8, 138-142.
- VON EIFF, A. W. Ueber die Möglichkeit einer Grundumsatzsenkung durch Psychische Beeinflussung. *Artzt.-Forsch.*, 1950, 4, 611.
- VON KRAFFT-EBING, R. *Eine experimentelle Studie auf dem Gebiete des Hypnotismus*. (2nd ed.) Stuttgart: 1889.
- WEISS, B. Electrocardiographic indices of emotional stress. *Amer. J. Psychiat.*, 1956, 113, 348-351.
- WEITZENHOFFER, A. M. *Hypnotism: An objective study in suggestibility*. New York: Wiley, 1953.
- WELLS, W. R. Experiments in waking hypnosis for instructional purposes. *J. abnorm. soc. Psychol.*, 1924, 18, 389-404.
- WELLS, W. R. The hypnotic treatment of the major symptoms of hysteria: A case study. *J. Psychol.*, 1944, 77, 269-297.
- WEST, H. F., & SAVAGE, W. E. Voluntary acceleration of the heart beat. *Arch. intern. Med.*, 1918, 22, 290-295.
- WEST, L. J. Psychophysiology of hypnosis. *J. Amer. Med. Ass.*, 1960, 172, 672-675.
- WHITEHORN, J. C., LUNDHOLM, H., FOX, E. L., & BENEDICT, F. G. Metabolic rate in "hypnotic sleep." *New England J. Med.*, 1932, 206, 777-781.
- WOLF, S., & WOLFF, H. G. *Human gastric function: An experimental study of a man and his stomach*. (2nd ed.) New York: Oxford Univ. Press, 1947.
- WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology*. (Rev. ed.) New York: Holt, 1954.
- YEAGER, C. L., & LARSON, A. L. A study of alpha desynchronization in the electroencephalogram utilizing hypnosis. Paper presented at American Electroencephalographic Society, Santa Fe, October 1957.
- ZWICK, C. G. Hygiogenesis of warts disappearing without topical medication. *Arch. Dermatol. Syphilol.*, NY, 1932, 25, 508-521.

(Received May 19, 1960)

THE NUMBER CONCEPT: A PHYLOGENETIC REVIEW

FRANK WESLEY

Portland State College

The data concerning number concepts of animals which have been reported so far do not agree with the general relationship between position in the phylogenetic scale and behavior. Rensch and Altevogt (1953) working with an elephant and Hicks (1956) with monkeys had limited success in establishing a "threeness" concept, while Koehler (1943) and his collaborators (Arndt, 1939; Braun, 1952; Lögler, 1959; Marold, 1939; Sauter, 1952; Schiemann, 1939) obtained a "sevenness" level on several species of birds.

Salman (1943) reviewed the number capacities of animals and found inadequate controls which allowed operation of rhythmic cues in most studies reported prior to 1939. Salman and other reviewers (Honigman, 1942; Koehler, 1951; Thorpe, 1956) considered rhythmic cues and other extraneous variables very well controlled in the studies reported in 1939 and thereafter, but did not mention the omission of certain operating procedures considered standard in the United States literature. The bird studies, originally reported in German, have been uncritically accepted by Hicks (1956), Morgan (1956), Newman (1956), and other American writers. It will therefore be the purpose of this paper to reexamine in detail the methodology used in studies reported since 1939.

Definition

There is considerable agreement among most investigators in the definition of a number concept. An

animal is usually required to solve a problem without the aid of immediate physical variables. External cues such as size, shape, color, brightness, tactile, odor, etc., as well as internal ones arising from rhythmic motor patterns or other visceral or kinesthetic feedback should either be absent or randomized from trial to trial, so that the numerosity of the stimulus constitutes the only constant variable. Experiments which were designed to include any of the above physical variables on a non-randomized basis will be omitted from this review. Included, however, will be those in which a number concept is reported by the authors though immediate and constant cues could have been responsible for the observed behavior.

BIRDS

Arndt (1939) tested number ability with various tasks in nine pigeons which received an average of 4,000 trials. To prevent rhythmic sequences Arndt presented peas on a turntable, exposing one pea at a time. Delays from pea to pea were from 1 to 60 seconds. His pigeon "Blaugrau" mastered the pecking of five peas only and would not touch a sixth pea when it appeared in the open slot. Another pigeon "Grau" learned a fourness problem on its first trial. In a tube experiment Arndt dropped peas from behind a screen at intervals varying from 1 to 20 seconds. On the animal's side the pea fell into a cup-shaped receptacle at the end of the tube. With this method pigeon

"Braunweiss" responded correctly with 55% in a twoness problem during the last hundred of 915 trials. When subsequently trained for a threeness problem, it responded correctly on the first seven trials, which means that without any negative transfer from the previous problem it picked the now correct third pea. Such an initial and highly accurate response strongly suggests the presence of extraneous cues. Arndt, however, looked upon it as "progress in learning" not realizing that even the most optimal "learning to learn" situation requires some negative transfer. In another experimental arrangement Arndt employed lid-covered boxes on a turntable. Again, only one box appeared in an open slot at any one time. With a twoness task one wheat kernel or one pea was placed into each of two successively appearing boxes. One pigeon "Blauweiss" learned to open these two successive boxes, but would not open a third box. When the two baits were distributed within three boxes, (1, 0, 1) the second box being empty, "Blauweiss" exhibited immediate learning, opening now three boxes, and leaving the fourth one untouched. From this behavior Arndt concluded that the bird had not learned to open a certain number of boxes, but learned to eat a certain number of peas. Gradually, within 6,000 further trials it learned to take six peas out of six boxes, not opening the seventh box. During the above experiments Arndt noticed that the birds would usually remain at the slot of the apparatus after they had responded correctly, and would turn away only after the turntable turned to present the negative stimuli. Arndt tested the possibility of differential acceleration as an extraneous cue, which may have been possible, since the turn-

table was operated manually. He asked another experimenter to turn the table for 600 trials and observed no differential results when compared with data from his own manipulations. He failed to note that the other experimenter also knew the correct number, and that subjective acceleration cues may have remained constant from one experimenter to another. In his review Thorpe (1956) describes Arndt, among other experimenters, as having "adopted quite extraordinary precautions to avoid errors of the 'Clever Hans' type" (p. 344). But "Clever Hans" could also solve problems when given by another experimenter who knew the correct answer.

Arndt obtained 65% correct responses as the highest level of performance on sets of 100 trials during thousands of trials. Such a low level of mastery and the frequent absence of negative transfer do suggest extraneous cues with both the tube and turntable experiments. Auditory variables which could have arisen from the experimenter or from a rattling of baits in the boxes during the turning were not controlled. Olfaction received no attention and boxes were not baited beyond the desired number. Another extraneous variable in Arndt's methodology could have been the nonrandomization of the amount of food ingested. Since it can be assumed that most of the peas were of equal size, visceral feedback could have presented a constant and immediate stimulus, and the correct response could have been based on quantity of food rather than on a mediated numerical concept. Thus, a quantity of food, an odor, a noise, or a "subjective" turning speed may alone or in combination be responsible for the results observed. Arndt's methodology therefore, does not war-

rant the conclusion that behavior based on numerosity was exhibited.

Concurrently with Arndt, Marold (1939) tested several parakeets on simultaneous and successive tasks. One parakeet was trained to discriminate between groups of two and three kernels. No learning was exhibited within 500 trials. The bird, apparently, depended too strongly on figure aid and changed to a position habit whenever the figure aid was withdrawn. To break this position habit Marold allowed the bird to eat the negative group of kernels after a positive response and observed positive results on the sixth block of 100 trials. The correctness level, however, did not rise above 57% correct within 1,100 trials. On the successive task Marold used rows of kernels and required her birds to eat x kernels without touching the $x+1$ kernel. Marold's parakeet "Grün" was trained to eat two kernels from a row varying from three to seven kernels. The distance between the kernels was altered from .5 to 0 centimeter and with decreasing distance, decreasing accuracy was observed. At the end of 900 trials the bird responded 87% correct, but the percentage dropped to 44 on a subsequent block of 100 trials which involved a further decrease in spacing. In an additional block behavior resembling experimental neurosis was reported.

Throughout her experiments Marold reported large individual differences, but she concluded that these differences arose from individual differences in treatment and in "Einfühlung." Such a statement suggests that the birds did receive differential treatment intentionally or unintentionally, which may have accounted for some of the results observed. Marold's simultaneous

discrimination task was not free from differential size cues. Likewise, extraneous distance cues were present during her successive task, resulting in experimental neurosis at zero distance. Such behavior resembles the inability to differentiate between cues immediately present in the environment. The presence of these and other extraneous cues makes it difficult to ascribe Marold's observation to numerical behavior alone.

Schieman (1939) used a new method for investigating the ability of birds to act successively to numbers. He confronted his jackdaws with a row of 10 covered dishes and required them to uncover their lids in sequence. Baits were differentially distributed according to a prearranged pattern, so that sometimes a dish would contain two or more baits and sometimes none. His birds had to uncover a different quantity of dishes to obtain x baits. Odor cues were not controlled since the dishes beyond the correct number which were not to be uncovered did not contain bait. One jackdaw "Blau" exhibited its upper limit of $x=6$ and performed with 65% correct during 886 trials. Another jackdaw learned within 1,000 trials to differentiate between the eating of two, three, four, and five baits.

Schiemann reported that performance was lower on a task which required the opening of x dishes rather than the eating of x baits. Schieman did not vary the amount of food per trial, so that on any one number task this could have presented immediate cues from visceral feedback. Such a hypothesis could explain the high performance with "baits eaten" and the random performance with "dishes uncovered."

Schiemann attempted further a combination of successive and simul-

taneous number discrimination. He presented his jackdaw "Grün" a stimulus card containing either two or four dots. According to this sample-number "Grün" was subsequently required to peck the indicated number of baits from a plate. After 1,200 trials a 70% correct response was obtained in the last 100 block. Schiemann believed that this demonstrated a success in the ability to act out a previously seen number. It should, however, be noted that the size of the sample dots remained constant throughout this task. With stimuli differing in one physical dimension this task may be compared with the disjunctive RT experiments so need not be related to number concepts. If unknown samples were presented, Schiemann states, the jackdaws appeared to be "completely helpless."

Koehler (1943) worked intensively with a 9-year-old raven named "Jacob" which received a total of approximately 12,000 trials during 794 working hours. On a series of trials "Jacob" learned to discriminate successfully between piles of baits having the following ratios: 4:5, 4:6, 6:5, and 7:6 (the first number indicating the positive stimulus). After having mastered these tasks it was not possible to train "Jacob" to discriminate on a 5:6 problem, though several hundred trials were administered with and without punishment and interspersed with rest periods. A naive bird was likewise unsuccessful. Koehler noted in a later film of the experiment that his assistant had a tendency to place the positive group closer to the forward margin of the experimental board during the discrimination series. Koehler did not mention whether correction of this placement cue preceded the 5:6 problem, but if it did the failure

could be explained. As in the famous case of Kinnebrook the assistant was replaced but manual placing of baits was continued.

The most difficult task which "Jacob" learned was a multiple choice task with a sample-indicator. A sample card was placed on the ground indicating the required number by means of irregular dots. Around it five covered dishes were placed with their lids showing irregular dots from numbers two to six. The dots on the positive lids differed in size, shape, and configuration from the ones on the sample. "Jacob" was able to obtain the reward of either grain, fruit, cheese, or meat when all aids were withdrawn and when the dots were replaced by irregular pieces of plasticine. The breaking and kneading of the plastic material as well as placing it on the lids was done manually, and again inadvertent cues arising from this manipulation should not be excluded in the evaluation of the obtained results. Odor control was likewise seriously lacking and initiated only after Tinbergen reported to Koehler that a jay was able to detect mealworms by odor. During the 481 trials which were presented with the irregular plasticine dots only five were partially odor controlled by baiting several dishes. "Jacob" responded correctly on all of these five trials.

Braun (1952) worked with three parrots to investigate some combination tasks. For positive reinforcement hempseed or cheese was used. Negative reinforcement consisted of punishment with a stick but was applied only when "absolutely necessary." At other instances during her experiments Braun made loud noises, threw a wet sponge, or pulled tail feathers as methods of negative reinforcement.

One of Braun's parrots performed on a dish-row problem during a six-ness task with an average of 75% correct responses. It is difficult, however, to evaluate the results and ascribe them only to a numerical concept since the physical proximity which the experimenter must have maintained with her animals in order to administer the various methods of punishment could have presented a host of extraneous cues. Furthermore variables such as odor and amount of food ingested were not controlled.

Eight magpies were used as subjects by Sauter (1952) who repeated some of the above tasks. Her food dishes were baited and covered in the observation room and manually spaced 10-12 centimeters in the experimental room. Her rewards were a variety of foods, with type and amount remaining equal within one task. A scare apparatus which was rarely used served for negative reinforcement. She tested four magpies on a dish-row. X baits were distributed in 25 different ways into 10 dishes. Magpie "Prinz" accomplished an $x=3$ problem on this task, but showed no transfer effect when a row of ten mealworms replaced the row of 10 dishes. The upper limit, $x=7$, was reported to have been reached with magpie "Felix" performing at a level of 74% correct at the end of 100 trials.

Half of Sauter's bait distributions on the dish-row problem did not contain zero spacings and if odor cues, which were not controlled, were postulated, they could account for 50% of the correct responses. An additional 25% correct responses could be assumed if chance behavior occurred on 50% of the trials which did have zero spacings and therewith a possible odor control. Thus, the odor variable alone could explain the

74% correct level reported for "Felix" during the $x=7$ task. Odor remained likewise uncontrolled in Sauter's simultaneous experiments. Dishes were manually placed and the selection and production of the irregular pieces of plasticine did not follow a prescribed procedure to assure the nonoccurrence of the experimenter's unintentional cues.

Braun's parrot "Jako" was again used by Lögler (1959) to perform on 16,076 additional trials on various number combinations. One such task involved successive presentations of flashes of light which varied in number and served to indicate the correct number of baits to be chosen on a dish-row. Numbers up to seven could be acted out in this way. Stimulus generalization of numerosity was reported on lower numbers when the visual indicators were replaced by auditory ones. On a single dish-row problem "Jako" was successful in obtaining eight baits which were differentially distributed into 11 dishes. The bird reached significant results after performing on chance level for 600 trials. Though odor was not controlled its possible interaction on the successive problem is not likely since 4/5 of Lögler's bait distributions contained one or several zero spacings. But, again, the manual placing of all dishes, the nonrandomization of food quantities, and the occasional deviations from the intended methods of scaring could have contributed extraneous cues.

All of the above reported bird studies used variable and highly subjective types of negative reinforcement. Most experimenters designed a scare apparatus intended for uniform punishment, but abandoned it early in their experiments. Arndt (1939) changed over to bait withdrawal while Marold (1939) used

blasts of air to blow away negative kernels but allowed them to be eaten in other instances. She also reported spraying of water and gypsum powder into the birds' faces. At times Schiemann (1939) "scared" the birds but darkened the room at other occasions. It is not clear in most of the above reported experiments how often and in which instances the various punishment methods were used. Honigman (1942) and Salzman (1943) who reviewed the above experiments believed them well controlled and Thorpe (1956) more recently termed them "technically beyond reproach" (p. 349) and stated that the use of the punishment apparatus made it impossible for the experimenter to give inadvertent signs. He does not mention the abandonment of the apparatus and the substitute method reported by the various authors.

Another serious lack common to the above bird studies is the absence of odor controls. Only Koehler (1943) controlled it partially in 37 trials out of an approximate total of 55,000 trials given by the above experimenters who offered such variable baits as flour, bread, seeds, fruits, cooked and raw meats, cheese, and others. The general assumption that birds are insensitive to odor may be quite fallacious. In a well controlled experiment Zahn (1933) found odor sensitivity equal and surpassing human thresholds on five different odors in experiments with pigeons, blackbirds, blue titmouse, robins, and hedgesparrows.

Manual placing of turntables, cups, lids, baits, dots, or plasticine was also present in all of the above bird studies. It was reported by Arndt (1939) and Koehler (1943) to have influenced their results on certain occasions. Elimination of these and

other cues could have been assured only by complete mechanical presentation of the stimulus components, a methodology not adopted by any of the experimenters.

FISH

The counting capacities of minnows, stickleback, and other small fish were investigated by Rossmann (1959) who found that innate preferences of bait size, stimulus density, and motility interfered with numerosity throughout prolonged training on the simultaneous discrimination task. One motor act and one tonal quantity, however, could eventually be differentiated from two acts or two tones on the successive task. One minnow, e.g., required a sequence of 170 negative reinforcement trials before it learned to eat the first bait without touching the second. The experiment was well controlled in regard to odor, bait-size, and rhythm. Training to numbers above one could not be established and Rossmann concluded that a number concept in fish can therefore not be postulated.

MAMMALS

Rodents

Hassmann (1952) experimented with 13 squirrels employing the methodology used by Koehler and his collaborators with the bird subjects. She used a variety of nuts and seeds as reward and scaring with a broom as negative reinforcement. On the dish-row task "Grauhörnchen" was reported to have demonstrated a concept of fiveness and "Hans" one of sixness. Hassmann's simultaneous task required the differentiation between five lid-covered dishes each bearing a number from three to seven. These numbers were indicated by irregular dots that

changed in size and position from trial to trial. One squirrel "Hexer" could differentiate the seven-lid from the three, four, five, and six-lids.

There are several observations by Hassmann which strongly suggest the presence of extraneous cues. A new fourness task was solved with an initial correctness equal to a previous threeness task involving 600 trials. Hassmann did not interpret this as a possible indicator of extraneous cues, but termed this behavior "a surprising success in learning." Odor was not controlled since negative dishes were not baited during a total of approximately 15,000 trials, in spite of the fact that Hassmann reported an aversion on the part of one of her animals from newly painted dishes (Hassmann, 1952, p. 299). On one occasion an unplanned odor control was reported. A squirrel pushed a positive lid aside, without "seeing" the peanut in it. It went to some negative dishes but returned later to the positive dish, opened it completely and obtained the bait. If "Peter" had smelled the peanut, Hassmann maintains, it would have continued to displace the positive lid on its initial attempt. Hassmann did not include peanuts in her previous list of rewards and it is difficult to determine the amount of acquaintance "Peter" had with this type of reward and its odor. The performance was very much like that which Tinklepaugh (1932) observed in monkeys when rewards were changed during a delayed reaction test.

Aside from odor the manual placing of all cups, lids, and sample dots could have presented additional extraneous cues. Hassmann's methodology should be scrutinized since she reported the successful learning of an oddity task in which the cue for solution was always numerosity. If

this performance occurred without the aid of extraneous cues it would represent one of the highest conceptual achievements on subprimate level.

Wesley (1959) investigated numerosity in the rat in a successive task in which the animals were required to enter a "second" open alley without previously entering a "first" open one. The alleys, their location, and their total number changed from trial to trial. Some significant runs were obtained only by massing trials at the end of daily practice sessions, linked with nonreinforcement after an initial prolonged corrective training. Osgood¹ pointed out that it is possible the animal responded by avoiding the first open, negative door rather than by entering the second open one. Thus as in the case of Rossmann's fish the rats may have responded only to oneness.

The rats' capacity to discriminate by numerosity was further investigated by Wesley on a multiple serial visual discrimination apparatus. Rats were able to perform on a twoness task after approximately 100 trials and showed negative transfer to a subsequent threeness task. Discrimination of threeness was acquired but not maintained after the exclusion of triangularity.

Elephant

The visual learning capacity of an elephant was studied by Rensch and Altevogt (1953) who presented three- and four-dot patterns on stimulus cards. After almost 100 trials the elephant was able to distinguish correctly between irregular dots on a 3:4 discrimination problem, but only with constant arrangement of the

¹ C. E. Osgood, personal communication, October 1960.

stimulus dots. Since the positive three-dot pattern was always presented in one of seven arrangements and the negative four-dot pattern in one of five arrangements, the elephant could have solved the entire task by learning five different Gestalten, and it is therefore questionable whether the animal had the abstractive capacity the experimenters suggest.

Monkeys

Douglas and Whitty (1941) reviewed the literature of number appreciation in subhuman primates and tested four baboons in a visual discrimination experiment. They presented either one or two successive flashes and required a different response to each cue. When subsequently they were equated for duration the proportion of correct responses fell to a low value.

Kühn (1953) investigated the ability to differentiate visually between black dots of varied sizes and arrangements. His 2-year-old rhesus monkey "Lola" received a total of 18,718 trials within 439 working hours. Kühn used 50 discrimination cards per number throughout his experiment and presented these on training and on test trials. He reported learning to discriminate number on an 8:6 task, but it should be noted that the cards of the six series were presented 500 times prior to this task, always designating the negative stimulus. Responses may have occurred to individual cards and not necessarily by means of the number concept they presented. A similar type of learning may have been involved in the solution of the 8:7 task.

Hicks (1956) investigated the number concept in eight adolescent rhesus monkeys. His methodology was free of extraneous cues, since he

introduced new and different stimulus cards during test trials. All of his animals performed above chance on a threeness problem, though some with rather moderate proficiency. Hicks compares his results with other studies and assumed that the 8:7 discrimination level observed by Kühn represents a true number concept but he had some doubt whether his own positive results indicated a number concept *per se*, since in all tests of number concepts the stimuli possess other characteristics than number. If, however, such a definition is employed no number concept *per se* could ever be demonstrated even on a human level, as stimulation always involves physical characteristics in addition to numbers. Heidebreder (1946), e.g., could not present twoness to her human subjects without involving objects, gestalten or size.

CONCLUSIONS

The performance of birds on a sevenness level has been compared to the human level of subitizing, an estimating of number without counting, where seven seems to form the average upper limit. (Jevons, 1871; Miller, 1956.) A re-examination of the methodology of these bird studies, however, makes such a comparison invalid and questions performance at any numerical level. Phylogenetically, the monkey would be expected to perform closer to the human level, but at present threeness is the only level unequivocally established with this species. The numerical capacity above threeness needs further investigation with monkeys, as numerosity in general needs to be studied further throughout the entire phylogenetic scale. To free future experiments from the influence of extraneous cues the presentation of the stimuli should be

mechanical and should randomize time, distance, size, and amount of food and should control odor, noise, and other possible immediate cues. It is very likely that the use of rigid

experimental controls will show that performance involving number concepts is congruent with the phylogeny of behavior as observed in other types of tasks.

REFERENCES

- ARNDT, W. Abschliessende Versuche zur Frage des Zahlvermögens der Haustaube. *Z. Tierpsychol.*, 1939, 3, 88-142.
- BRAUN, HILDEGARD. Über das Unterscheidungsvermögen unbenannter Anzahlen bei Papageien. *Z. Tierpsychol.*, 1952, 9, 40-91.
- DOUGLAS, J. W. B., & WHITTY, C. W. M. An investigation of number appreciation in some sub-human primates. *J. comp. physiol. Psychol.*, 1941, 31, 129-142.
- HASSMANN, M. Von Erlernen unbenannter Anzahlen bei Eichhörnchen. *Z. Tierpsychol.*, 1952, 9, 294-321.
- HEIDBREDER, E. The attainment of concepts: I. Terminology and methodology. *J. genet. Psychol.*, 1946, 35, 173-189.
- HICKS, L. H. An analysis of number concept formation in the rhesus monkey. *J. comp. physiol. Psychol.*, 1956, 49, 212-218.
- HONIGMAN, H. The number concept in animal psychology. *Biol. Rev.*, 1942, 17, 315-337.
- JEVONS, W. S. The power of numerical discrimination. *Nature, Lond.*, 1871, 3, 281-283.
- KOEHLER, O. "Zähl"-versuche an einem Kolkrahen und Vergleichsversuche an Menschen. *Z. Tierpsychol.*, 1943, 5, 575-712.
- KOEHLER, O. The ability of birds to count. *Bull. anim. Behav.*, 1951, No. 9, 41-45.
- KÜHN, E. Simultanvergleich gesehener Mengen beim Rhesusaffen. *Z. Tierpsychol.*, 1953, 10, 268-296.
- LÖGLER, P. Versuche zur Frage des Zählvermögens an einem Graupapagei und Vergleichsversuche an Menschen. *Z. Tierpsychol.*, 1959, 16, 179-217.
- MAROLD, E. Versuche an Wellensittichen zur Frage des "Zähl"-vermögens. *Z. Tierpsychol.*, 1939, 3, 170-223.
- MILLER, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, 1956, 63, 81-97.
- MORGAN, C. T. *Introduction to psychology*. New York: McGraw-Hill, 1956.
- NEWMAN, J. R. *The world of mathematics*. Vol. 1. New York: Simon & Schuster, 1956.
- RENSCH, B., & ALTEVOGT, R. Visuelles Lernvermögen eines indischen Elefanten. *Z. Tierpsychol.*, 1953, 10, 119-134.
- ROSSMANN, ANGELIA. Über das "Zähl"-Vermögen der Fische. *Z. Tierpsychol.*, 1959, 16, 1-18.
- SALMAN, D. H. Note on the number conception in animal psychology. *Brit. J. Psychol.*, 1943, 33, 209-219.
- SAUTER, ULRIKE. Versuche zur Frage des "Zähl"-Vermögens bei Elstern. *Z. Tierpsychol.*, 1952, 9, 252-289.
- SCHIEHMANN, K. Vom Erlernen unbenannter Anzahlen bei Dohlen. *Z. Tierpsychol.*, 1939, 3, 292-347.
- THORPE, W. H. *Learning and instinct in animals*. Cambridge: Harvard Univ. Press, 1956.
- TINKELPAUGH, D. L. Multiple delayed reaction with chimpanzees and monkeys. *J. comp. physiol. Psychol.*, 1932, 13, 207-243.
- WESLEY, F. Number concept formation in the rat. *Z. Tierpsychol.*, 1959, 16, 605-627.
- ZAHN, W. Über den Geruchssinn einiger Vögel. *z. vergl. Physiol.*, 1933, 19, 785-796.

(Received June 7, 1960)

HYPNOTIC AGE REGRESSION

AUBREY J. YATES¹

University of New England, Australia

The alleged phenomenon of hypnotic age regression refers to the apparent fact that a subject (*S*) who is told under hypnosis that he is, e.g., 4 years old, may behave in a manner which is characteristic either of his own behavior at that age, or of children in general at that age. It should be noted carefully that the phenomenon may refer to the reactivation of behavioral characteristics of *S* himself; or may refer to a more general revivification of childlike behavior.

According to Platonow (1933), hypnotic age regression was first demonstrated clinically in 1893 by Kraft-Ebing. In spite of a good deal of clinical interest in the alleged phenomenon, experimental interest in the problem remained dormant until the publication of Platonow's report, in which he claimed that hypnotic age regression had been objectively demonstrated in three *Ss*, using the Binet test, and that the general behavior of the *Ss* in the regressed state showed characteristic childlike features. This conclusion was challenged by Young (1940), who claimed that *Ss* were able in the waking state to reproduce the test and general behavior of a young child voluntarily, and with greater accuracy than hypnotized *Ss* who had been regressed. Curiously, this challenging problem of the authenticity of hypnotic age regression has not received a great deal of attention since Sears (1943) first reviewed it briefly. A review of the literature

suggests, however, that the problem is now much more clearly defined and a number of facts can be accepted as reasonably well-established. The principal controversies have centered around the disputes as to whether regression (partial or complete) can be demonstrated; if so, whether the regressed state can be simulated by *S* or represents a genuine reactivation of previous habit-systems or personality organizations; and finally, whether or not hypnosis is an essential part of the process.

This review will cover the types of measures which have been used to compare the test performance and behavior of the *S* in the waking state with that in the hypnotic state, the various conditions under which the performance is recorded, the principal established results, the main theories which attempt to account for the phenomenon, and the methodological problems involved. Some suggestions for future research will be made.

TYPES OF COMPARISON

A distinction may be drawn between direct and indirect comparisons² of the waking and regressed states; and between the use of measures which are susceptible to simulation to a greater or less degree, and those which on the whole are not susceptible to simulation.³

² The problem is similar to that faced by the clinical psychologist attempting to measure deterioration (Yates, 1956).

³ The distinction is, of course, an arbitrary one, but it does help in organizing the field.

¹ Now at the University of Western Australia. Acknowledgement is due to D. Howie for valuable criticisms of this paper.

Direct Comparisons

A few studies have directly compared *S*'s present regressed performance with his known performance on the same measure at a time corresponding to the regressed age level.

Simulable Measures. Sarbin (1950b) compared the performance of 12 *S*s regressed to the age of 8 years with their performance on the same test when they were actually 8 years old. Similarly, Orne (1951) was able to compare the original and regressed drawings of one of his *S*s.

Nonsimulable Measures. True (1949) regressed his *S*s successively to ages 10, 7, and 4 years, and asked them what day of the week their birthday and Christmas Day fell on in that year. Best and Michaels (1954), and Reiff and Scheerer (1959) performed similar experiments, while the latter additionally attempted to reactivate factual information about childhood experiences (e.g., names of teachers and classmates) which *S*s claimed to be unable to recall in the waking state.

These studies appear to represent the only sources of direct comparison thus far made.

Indirect Comparisons

In the majority of studies on hypnotic age regression the present regressed performance of *S* is compared with the average known performance of normal *S*s of the age to which regression is induced.

Simulable Measures. There are three sources of evidence here. In the field of mental testing, investigators have used the Binet test (Keir, 1945; Platonow, 1933; Spiegel, Shor, & Fishman, 1945), the Wechsler-Bellevue Intelligence Scale (Kline, 1951), the Otis Performance

tests (Kline, 1950), the Word Association Test (Dittborn, 1951; Keir, 1945), and various kinds of motor tasks, such as drawing a man, handwriting, etc. (Orne, 1951; Platonow, 1933). Second, the general behavior of *S* in the regressed state has been compared with that of normal children of that age-level (Keir, 1945; McCranie, Crasilneck, & Teter, 1955; Orne, 1951; Platonow, 1933). Third, *S* has been placed (in the regressed state) in situations known to evoke intense fear responses in many young children and his behavior observed (Kline, 1953a).

Nonsimulable Measures. Three kinds of information have been utilized. In the field of mental testing, the Bender Gestalt Test has been used (Crasilneck & Michael, 1957); and the Rorschach test (Bergman, Graham, & Leavitt, 1947; Keir, 1945; Mercer & Gibson, 1950; Norgar, 1952; Orne, 1951), the argument, of course, being that it would be difficult for adults to simulate the performance of young children on these tests, especially in the case of the unstructured Rorschach test. Second, a number of physiological measures have been recorded while *S* was in the regressed state. These include the presence or absence of the plantar response (Gidro-Frank & Bowersbuch, 1948; True & Stephenson, 1951) and the Babinski reflex (McCranie et al., 1955); changes in indices such as blood pressure, pulse, and respiration rates, and psychogalvanic reflex (Kline, 1960; True & Stephenson, 1951); and changes in EEG characteristics (McCranie et al., 1955; True & Stephenson, 1951), the argument in the latter case being that there are characteristic differences between the records of adults and children.

Third, special mention must be made of the advances in technique recently reported by Reiff and Scheerer (1959). In accordance with Scheerer's general theoretical position, they abandoned the use of mental age tests, and utilized instead the notion of developmental levels, analyzing *S's approach* to the solution of various problems, rather than his correct or incorrect *responses*. They used a number of ingenious tests which they consider it would be particularly difficult to simulate. Thus, in the Lollipops test, the regressed adult was given a lollipop after making mud-pies, while his hands were still dirty. If true regression had taken place, the child regressed to 4 years would, they argued, naturally not worry about his dirty hands when accepting the lolly; the adult simulating regression would do so. On the cognitive side, they used a Pledge of Allegiance test (writing the pledge after reciting it), a Clock test (telling the time), a Left and Right test (identifying left and right e.g., in persons sitting opposite), an Arithmetic test; Piaget's Hollow Tube Test (identifying the order in which colored beads will emerge from a hollow opaque tube after it has been rotated), and a Word Association test. On all of these tests, children show characteristic changes in modes of response with increasing age. On the Word Association test, for example, the most popular responses with children are quite different from those found in adults. They further argued that the adult who simulated a child's responses on this test (correctly or incorrectly) would show increased reaction time, since he would first need to inhibit his natural adult response tendency.

TESTING CONDITIONS

Six distinct testing conditions have been employed. Four of these may be regarded as control conditions for the two extreme conditions of performance in the normal waking state compared with performance in the suggested regressed state under hypnosis. The six conditions are:

1. Normal waking state—With some exceptions (e.g., Platonow, 1933; Sarbin, 1950b; Spiegel et al., 1945), nearly all investigators record performance in this condition.

2. Normal waking state, with deliberate (not simulated) attempted recall of earlier events—This control condition has seldom been used, though it is clearly essential to the validity of results such as those obtained in True's (1949) experiment.

3. Normal waking state, with instructions to simulate regression to a particular age-level—This control condition was used by Reiff and Scheerer (1959).

4. Standard hypnotic state—This represents a control for the effects of hypnosis per se, and has rarely been used.

5. Hypnotic state, with instructions to simulate regression to a particular age-level—This condition has been used only once experimentally (Crasilneck & Michael, 1957).

6. Hypnotic state, with direct suggestion by *E* that *S* is now a certain age.

No single study has used all of these conditions, and only one (Crasilneck & Michael, 1957) has used as many as four.

It would be an important requirement of any experiment in this field that the judgments of behavior in the various conditions should be made in ignorance of the particular condition

in which *S* is placed at the time of assessment, i.e., the double-blind technique should be used.

PRINCIPAL RESULTS

Regression can be produced under hypnosis (Condition 6), the extent and accuracy of the regressed behavior being a matter of considerable dispute. Thus, Sarbin (1950b) using direct comparisons of Binet performances, found that under hypnosis not one of his nine hypnotizable *Ss* achieved a mental age as low as that on the original test occasion. Using indirect comparisons, Crasilneck and Michael (1957) regressed their *Ss* to the age of 4, but found that the Bender Gestalt drawings were rated by independent judges as comparable with those of 7-year-old children. It is not surprising, therefore, that inconclusive results have been generally reported for nonsimulable complex tests such as the Rorschach, except for obvious measures such as number of responses and form-quality (Bergman et al., 1947; Orne, 1951). On the other hand, it is important to notice that such regression as is achieved is often remarkably successful. Kline (1950) showed that in spite of a very significant decline in score on the Otis Performance tests under regression (from 59.2 in the waking state to 24.5 under regression to 8 years), the IQ at the regressed ages showed less variability than is normally found when the test is repeated on separate occasions.

Reiff and Scheerer (1959) reported almost uniformly perfect regression under hypnosis on all of the measures they used. Thus, on the Clock test, adults regressed to age 7 made errors characteristic of children of that age level, whereas the simulat-

ing controls did not make such errors.

Regression, however, does seem to become more accurate as the functions measured become more specific. True (1949), using 50 *Ss* in the experiment described earlier, found that when regressed to ages 10, 7, and 4, 92%, 84%, and 62% of *Ss*, respectively, correctly identified the day of their birthday; while 94%, 86%, and 87% of *Ss*, respectively, correctly identified the day on which Christmas Day fell. Best and Michaels (1954) found negative results in a similar experiment, but they used only five *Ss* and their procedure differed in important respects from that of True. Equally remarkable results were obtained by Reiff and Scheerer with the Word Association test. Thus, while adults uniformly respond to the word "man" with "women," children equally uniformly respond with the word "work." Their hypnotically regressed *Ss* responded with the characteristic child's response while simulating *Ss* continued to use the adult response word. McCranie et al. (1955) reported the reinstatement of the Babinski reflex in 3 of their 10 *Ss* when regressed to the age of 1 month; while Gidro-Frank and Bowersbuch (1948) found significant changes in the plantar response, which were accompanied by changes in peripheral chronaxie. McCranie et al. (1955) did not, however, observe any significant change in EEG records in the regressed state.

Moody (1946), Ford and Yeager (1948), and Erickson (1937), have all reported the reinstatement of disabilities (wheal marks, homonymous hemianopsia, attacks of unconsciousness) under hypnotic regression; the disabilities no longer being present in the waking state.

The general behavior of the regressed *S* has been frequently reported as becoming more childlike, (e.g., Reiff & Scheerer, 1959) even to the extent of the appearance of the sucking response and loss of speech when regression is induced to a very early age (McCranie et al., 1955). In many respects too, the behavior is appropriate, not merely to the age to which *S* has been regressed, but in relation to the environment as it was at that time.

Regression can be produced in the waking state by asking *S* to simulate the suggested age (Condition 3). Under these circumstances, Crasilneck and Michael (1957) showed that Bender Gestalt drawings will reflect the simulated age level less successfully than is the case with hypnotic regression; and this finding has been amply substantiated by Reiff and Scheerer (1959).

Regression can be produced in the hypnotic state by asking *S* to simulate the suggested age (Condition 5). Under this condition, Crasilneck and Michael (1957) showed that Bender Gestalt drawings will not be significantly different in quality from those produced by direct suggestion under hypnosis.

The results obtained by Crasilneck and Michael (1957) for the waking state, waking state with simulated regression, hypnotic state with simulated regression, and hypnotic state with induced regression, for the Bender Gestalt test indicated that regression was not complete under any of the conditions, that regression could be simulated, and that hypnosis facilitated the production of regressed behavior. Reiff and Scheerer (1959) did, however, find complete regression to appropriate developmental levels.

There is some indication that emotional regression can be induced. Kline (1953a) regressed a female *S* to the age of 3 years and placed her in situations frequently found to produce intense fear responses in young children (being left alone, entering a dark passage, seeing a strange person oddly dressed, sudden appearance of a live snake, sight of a headless doll, and presence of a live mouse). All but the first and last situations produced realistic fear reactions in the regressed, but not in the waking, state, including involuntary urination. Reiff and Scheerer (1959), analogously in a play situation, found less repugnance to eating with filthy hands under regression.

The mere induction of hypnosis itself does not produce regressed behavior in the normal *S* (Bergman et al., 1947; Kline, 1953a), though there are clinical reports of "spontaneous regression" under hypnosis (Gill, 1948; Keir, 1945; Schneck, 1955).

Sarbin (1950b) has reported a correlation of +0.91 between a regression index and degree of hypnotizability. His hypnotizable *Ss* were regressed under hypnosis, and later asked to simulate regression in the waking state. A regression index (RI) was computed for *S* according to the formula:

$$RI = \left[\frac{MA(\text{simulated regression})}{MA(\text{original test})} - \frac{MA(\text{hypnotic regression})}{MA(\text{original test})} \right] \times 100$$

We may conclude from this brief survey of results that a *prima facie* case appears to have been made out for the assertion that under some conditions certain adults behave in ways which are characteristically

those of children, although many of the details remain to be filled in.

THEORIES OF HYPNOTIC AGE REGRESSION

Three theories have been proposed in attempts to account for the above results.

Neurological Theories

Platonow (1933) explained regression in terms of what he called Pavlov's "true physiology of the brain," using especially the notion of words as conditioned stimuli producing physiological, biological, and psychological changes:

the suggestion of previous ages brings forth a real organic reproduction of the engrams, the formation of which belongs to the earlier periods of the individual's life (p. 205).

Regression is facilitated under hypnosis because the latter involves general inhibition of the cortex except for the area receiving auditory impulses. Under these conditions, the auditory stimulus (suggestion of regression) most readily activates the appropriate engrams. This theory would appear to be derived from the much older distinction between cortical and subcortical brain-processes, the latter mediating primitive responses. It is interesting to note that McCranie et al. (1955) assert that lesions of Brodman's Area 4 result in the restoration of the Babinski reflex in chimpanzees and man, while simultaneous bilateral ablation of Areas 4 and 6 produces infantile motor behavior in lower primates.

Kline (1953b, 1954) has proposed a "neuropsychological" theory, derived from the experimental observation that what he terms habit progression, as well as habit regression, can be demonstrated (Kline, 1951; Rubenstein & Newman, 1954). Thus, in one study by Kline (1951), a 22-

year-old woman was able, under hypnosis, to produce the typical Wechsler-Bellevue record of a 65-year-old woman, even to the extent of obtaining a characteristic Deterioration Index score for that age-group. Kline's theory (1953b) postulates that

the actual state involved in such activity is not regression, not progression, but a central state of perceptual release or disorientation which permits activity in any dimension or direction of time-space orientation (p. 26).

Under hypnosis, there occur what Kline (1953b) calls "directional alterations from a central process" (p. 25) and he lays particular stress on the importance of transference relationships between *S* and the hypnotist. He does not, however, regard the phenomenon as simply involving role-taking (see below) and his theory, in particular does not regard the evidence obtained from psychometric measures as crucial, though he does not deny their relevance. As at present formulated, Kline's theory would seem to be too general to be amenable to disproof.

Habit-Reactivation

Hypnotic age regression may be regarded as a special instance of instrumental act regression. In the latter, if *S* possesses alternative response patterns to a given stimulus, the stronger will normally occur. If, however, the stronger is prevented from occurring, then the inhibited response pattern will be reactivated. It is possible that, in some way as yet unknown, hypnotic suggestion of regression may inhibit current response patterns, and hence permit the reactivation of "forgotten" response patterns. Although this theory as presented here is very general, it is surprising that no consideration has yet been given by any worker in this

field to the relationship between hypnotic age regression and instrumental act regression. Contrariwise, it would be predicted that newly acquired conditioned responses would be lost in hypnotic age regression. McCranie and Crasilneck (1955) attempted to test this hypothesis by setting up voluntary hand withdrawal, and involuntary eyeblink, conditioned responses. Under hypnotic regression, the former disappeared, the latter did not. Similar results were reported by LeCron (1952).

Reiff and Scheerer (1959) have put forward a theory of hypnotic age regression which is derived from a general theory of remembering, but which could equally be regarded as a more general theory of habit reactivation than the instrumental act regression theory. According to them: the act of recall becomes also an act of contemporary reconstruction of the past event, to a large extent dependent upon the state of the person at the time of the recall (p. 15).

Such memories can take the form either of remembrances or of memoria. In the waking state, remembrances are memories with a conscious autobiographic index (i.e., experienced as "being in my past") and involve therefore personal continuity. Memoria are memories without a current autobiographic past reference (e.g., motor skills, vocabulary, etc.). Although the distinction is not absolute, remembrances are usually related to an experiential context, whereas memoria are related to an environmental context. Both kinds of remembering may arise voluntarily or involuntarily.

Hypnotic age regression

makes possible a reinstatement of the forgotten personal past either in the form of remembrances, or in the form of memoria and earlier ego apparatuses (p. 52). . . . [In general] since . . . memoria are without an experi-

ential auto-biographic index, the ego can more easily activate appropriate memoria than remembrances (p. 49) [but] the age-regressed subject may remember events with the experience of an auto-biographic index. However, here the reference point is no longer the *actual* present but that point in the autobiographic past to which the subject has been regressed (p. 52).

Remembrance or memoria reactivated in this way are always involuntary.

In hypnotic age regression, therefore, the attempt is made to reactivate the general environmental and experiential contexts of the *S* at the age regressed to. If these remembrances can be activated sufficiently strongly, then individual items of behavior may be reactivated in the form of memoria. It should be noted that in regression, memoria and remembrances are experienced as occurring here and now, whereas in the normal state they can both be referred to the past, though only remembrances will have autobiographical references.

Role-Playing

This viewpoint has been well expressed by the assertion that

we may formulate the concept of age regression by saying that the prevailing psychological condition enables the individual to take the role appropriate to the imagined world (Orne, 1951, p. 220).

The most important exponent of this theory is Sarbin (1950a; Sarbin & Farberow, 1952). Sarbin's theory holds that all human interaction involves both *S* and *E* indulging in role-playing. The validity of role-playing depends upon a number of factors, including the validity of *S*'s perception of the interaction situation, the aptitude of *S* for playing a particular role, and the current organization of the self. Thus, if age regression is to appear

the *S*'s perception of the child's role must have some veridical properties; there must be present evidence of the role-taking aptitude, and the assigned role must not be incongruent with the *S*'s current self-perceptions. (Sarbin & Farberow, 1952, p. 119.)

Sarbin's theory has resulted in a number of interesting predictions, for example, that an *S* whose self-organization is relatively undeveloped should show greater regression than an *S* whose self-organization is normal. Unfortunately, little concrete evidence in support of Sarbin's theory, in so far as it relates to age regression, has yet been produced.

METHODOLOGICAL PROBLEMS

Many methodological problems present themselves in connection with hypnotic age regression. They may be grouped into five areas.

Types of Control

Mention has already been made of some of the conditions under which testing has been carried out. Logically, at least three sets of factors can be varied, giving eight possible combinations of testing conditions. Thus, *S* may be tested in the hypnotic or waking state, the comparisons made may be direct or indirect, and regression or simulation may be attempted. Additionally, account must be taken of the effects on performance of hypnosis, and the waking state, per se.

Criteria for Regression

It is obvious that complete biological regression is impossible (the *S* in the regressed state does not, for example, diminish in stature). The critical question therefore becomes: "Does the hypnotically regressed adult perform as he *imagines* a child to function or is his regressive behavior a revival of memoria, i.e. of cer-

tain aspects of his previous functioning?" (Reiff & Scheerer, 1959, p. 83). It should be realized at the outset that the fact that hypnotic age regression may be incomplete does not in itself prove the validity of the role-playing theory; nor does the demonstration of successful role-playing in itself disprove the validity of regressive phenomena. The most satisfactory evidence for the validity of regression would involve the demonstration that *S* performed in the regressed state in a manner similar to his behavior at that age when a child (thus involving *direct* comparison) together with the demonstration that he was unable, as an adult, either under hypnosis or in the waking state, to simulate the appropriate behavior. Evidence of this kind has thus far been presented only in isolated cases. It is not here intended to deny, of course, that in most instances of hypnotic age regression, both true regression and role-playing may be simultaneously involved.

Hypnotic Technique

Several important points have commonly been neglected. Criteria for measuring the depth of trance are usually not reported in sufficient detail. The speed with which regression is induced is possibly a critical variable and may explain the failure of Best and Michaels (1954) to repeat the results of True (1949). It is probably important to reinstate the earlier period gradually, rather than suddenly. The role of the hypnotist is often neglected—it has been argued that true regression becomes more likely if the hypnotist transforms himself into some person familiar to *S* at the regressed age, or at least into a neutral figure. Reiff and Scheerer (1959) lay particular stress on the importance of the instructions

given to *S*, who should be regressed to a specific date (e.g., a birthday) and not merely to a particular year. Fluctuations in performance should be controlled as far as possible by instructing *S* not to deviate from his regressed age-level.

Selection of Ss

Very little attention has been paid to this important aspect of the problem. Reiff and Scheerer (1959) lay great stress on the difficulty of obtaining suitable *Ss* who must be relatively free from anxiety (severe anxiety about events happening at age 4, for example, might well lead to resistance to regression to that age), must be suitably motivated, and, of course, satisfactorily hypnotizable. Such *Ss* are relatively rare. Selection of control *Ss* has been even more neglected. Reiff and Scheerer, for example, give few details about their control *Ss* and do not seem to realize that the experimental and control *Ss* should have been carefully matched on all relevant variables (including hypnotizability).⁴ This failure was especially serious in that, with the exception of two measures, Reiff and Scheerer did not control for performance in the waking state, apparently assuming that all their *Ss* would perform normally.

Selection of Tests and Measures

The search for nonsimulable tests has been markedly improved by the suggestions of Reiff and Scheerer who, in addition to the tests they used, have made a number of ingenious suggestions for further research. The most significant measures thus far utilized are undoubtedly the Birthdays test of True (1949) and the

Word Association test used by Reiff and Scheerer (1954). The latter, however, prefer the use of developmental schedules to mental age scales, and are interested more in the process of solution than in the solution itself. While the distinction is not entirely academic, it certainly has not the importance attributed to it by Reiff and Scheerer. Two minor points may be noted: the tasks, measures, or developmental schedules used, should be appropriate to the age level regressed to; and care should be taken to prevent *S* from giving no response ("I don't know") except where such a response is explicitly predicted.

It may safely be said that no fully adequate experiment has been carried out in this field. Thus, the most recent study by Reiff and Scheerer (1959), although admirable in many respects, contained a number of serious faults: lack of control for performance in the waking state; failure to match experimental and control groups; and repeated testing at different age levels of the same *Ss* in the experimental, but not in the control group. Even more serious, it is clear from the description given of the experimental procedure, that the authors were aware, in the testing situation, of which *Ss* had been hypnotized, and which had not.

DISCUSSION

The potential importance of the phenomenon of hypnotic age regression can scarcely be overestimated. Apart altogether from its possible value in general psychotherapy (Kline, 1950), and its usefulness in particular for the treatment of war neuroses by regressing the patient to the traumatic situation and making him relive the experience, it does not seem to have been generally realized that the technique itself could

⁴ Their five hypnotized *Ss* were chosen from an original group of over 100 *Ss*.

provide a crucial test of the theory that learned responses are never "destroyed," but only supplanted and remain available for activation under appropriate circumstances. In light of this, it is surprising how little experimental work has been carried out in this area. Furthermore, a good deal of this work can hardly be said to attain even minimally acceptable levels of methodological adequacy.

Any acceptable theory of hypnotic age regression must take account of the apparent facts that regression can be simulated in the waking state; that the amount of regression is similar whether it is simulated under hypnosis, or suggested under hypnosis; and that regression becomes more "accurate" as the response regressed to becomes more specific. It seems likely that neither the role-taking, nor the habit reactivation theories, taken separately, will account satisfactorily for the observed facts. Thus, the role-taking theory is clearly embarrassed by findings such as those of True (1949) in relation to the recall of factual information under hypnosis, those of McCranie et al. (1955) in relation to the reactivation of physiological responses, and those of Reiff and Scheerer (1959) in relation to the Word Association test. It is probably necessary to recognize that both theories must be invoked, each accounting for some, but not all, of the facts. In this connection, it may be noted that a satisfactory explanation of the facts awaits the formulation of a valid *general* theory of behavior under hypnosis. Since, however, workers in this field are still struggling to elucidate basic concepts (Barber, 1958; Sutcliffe, 1960), it is probable that a more careful and

thorough examination of the *phenomena* encountered in hypnotic age regression will provide data highly relevant to this aim.

We may conclude, therefore, with some general suggestions concerning future research in this field. First, a crucial area of research is the problem of partial versus complete age regression. It seems clear that complete regression would be extremely unlikely in relation to complex items of behavior, since early habit-structures involving complex skills would surely be affected by subsequent growth of the skill, if only through the process of retroactive inhibition. On the other hand, relatively isolated items of knowledge (such as knowing on what day one's fourth birthday fell) might easily survive relatively unchanged by subsequent learning, to be reactivated under appropriate conditions. Mention has already been made of the necessity for a close examination of the conditions under which regression is induced.

Second, more attention should be paid to an analysis of simple aspects of behavior, rather than complex ones. For example, instead of using measures such as the Binet, attention could be concentrated on, e.g., developmental schedules, which objectively record the presence or absence of specific items of behavior at different age levels. The suggestions for research made by Reiff and Scheerer (1959) are particularly valuable in this connection. The use of conditioning techniques, as exemplified by the study of McCranie and Crasileck (1955) should also yield crucial information.

Third, much more attention should be paid to a careful description of the total behavior of *S* in the regressed state. Much has been made of the

fact that *S* behaves in a manner appropriate to his regressed age. Almost invariably, however, the description is highly selective. For example, as Orne (1951) has pointed out, regression implies also that all knowledge acquired subsequently to the age to which *S* has been regressed should be unavailable. In other words, *S* should no longer be cognizant of current affairs, political, social, or otherwise. It is extremely curious that no information of a concrete nature on this vital point is available, except for a few vague, general assertions.

Fourth, no attention has been paid to the study of the behavior of *S* in the regressed state over a substantial period of time. Practically all

investigators have restricted their observations to laboratory situations.

Fifth, the fact that hypnosis apparently facilitates simulated regression, but the addition of direct suggestion does not produce an improvement over hypnotically simulated regression requires further exploration. Thus far, evidence relating to this important point is restricted to results from a single study (Crasilneck & Michael, 1957).

The importance of the phenomena encountered in hypnotic age regression, and the advances in technique which characterize the investigations of Reiff and Scheerer (1959) should surely lead to a revival of interest in this problem.

REFERENCES

- BARBER, T. X. The concept of hypnosis. *J. Psychol.*, 1958, **45**, 115-131.
- BERGMAN, M. S., GRAHAM, H., & LEAVITT, H. C. Rorschach exploration of consecutive hypnotic chronological age level regressions. *Psychosom. Med.*, 1947, **9**, 20-28.
- BEST, H. L., & MICHAELS, R. M. Living out "future" experience under hypnosis. *Science*, 1954, **120**, 1077.
- CRASILNECK, H. B., & MICHAEL, C. M. Performance on the Bender under hypnotic age regression. *J. abnorm. soc. Psychol.*, 1957, **54**, 319-322.
- DITTBORN, J. Words associated to different age levels suggested under hypnosis. *Rev. Psiquiat., Santiago*, 1951, **16**, 105-107.
- ERICKSON, M. H. Development of apparent unconsciousness during hypnotic reliving of a traumatic experience. *Arch. Neurol. Psychiat., Chicago*, 1937, **38**, 1282-1288.
- FORD, L. F., & YEAGER, C. L. Changes in the electroencephalogram in subjects under hypnosis. *Dis. nerv. Syst.*, 1948, **9**, 190-192.
- GIDRO-FRANK, L., & BOWERSBUCH, M. K. A study of the plantar response in hypnotic age regression. *J. nerv. ment. Dis.*, 1948, **107**, 443-458.
- GILL, M. Spontaneous regression on the induction of hypnosis. *Bull. Menninger Clin.*, 1948, **12**, 41-48.
- KEIR, G. An experiment in mental testing under hypnosis. *J. ment. Sci.*, 1945, **91**, 346-352.
- KLINE, M. V. Hypnotic age regression and intelligence. *J. genet. Psychol.*, 1950, **77**, 129-132.
- KLINE, M. V. Hypnosis and age progression: A case report. *J. genet. Psychol.*, 1951, **78**, 195-206.
- KLINE, M. V. Childhood fears in relation to hypnotic age regression: A case report. *J. genet. Psychol.*, 1953, **82**, 137-142. (a)
- KLINE, M. V. Hypnotic retrogression: A neuropsychological theory of age regression and progression. *J. clin. exp. Hypnosis*, 1953, **1**, 21-28. (b)
- KLINE, M. V. Living out "future" experience under hypnosis. *Science*, 1954, **120**, 1076-1077.
- KLINE, M. V. Hypnotic age regression and psychotherapy: Clinical and theoretical observations. *Int. J. clin. exp. Hypnosis*, 1960, **8**, 17-42.
- LECROON, L. M. The loss during hypnotic age regression of an established conditioned reflex. *Psychiat. Quar.*, 1952, **26**, 657-662.
- MCCRANIE, E. J., & CRASILNECK, H. B. The conditioned reflex in hypnotic age regression. *J. clin. exp. Psychopathol.*, 1955, **16**, 120-123.
- MCCRANIE, E. J., CRASILNECK, H. B., &

- TETER, H. P. The electroencephalogram in hypnotic age regression. *Psychiat. Quart.*, 1955, 29, 85-88.
- MERCER, M., & GIBSON, R. W. Rorschach content in hypnosis: Chronological age level regression. *J. clin. Psychol.*, 1950, 6, 352-358.
- MOODY, R. L. Bodily changes during abreaction. *Lancet*, 1946, 2, 934-935.
- NORGARD, B. A. Rorschach psychodiagnosis in hypnotic regression. In L. M. LeCorn (Ed.), *Experimental hypnosis*. New York: Macmillan, 1952. Pp. 178-214.
- ORNE, M. T. The mechanisms of hypnotic age regression. *J. abnorm. soc. Psychol.*, 1951, 46, 213-225.
- PLATONOW, K. I. On the objective proof of the experimental personality age regression. *J. gen. Psychol.*, 1933, 9, 190-209.
- REIFF, R., & SCHEERER, M. *Memory and hypnotic age regression*. New York: International Univer. Press, 1959.
- RUBENSTEIN, R., & NEWMAN, R. The living out of "future" experiences under hypnosis. *Science*, 1954, 119, 472-473.
- SARBIN, T. R. Contributions to role-taking theory: I. Hypnotic behavior. *Psychol. Rev.*, 1950, 57, 255-270. (a)
- SARBIN, T. R. Mental age changes in experimental regression. *J. Pers.*, 1950, 19, 221-228. (b)
- SARBIN, T. R., & FARBEROW, N. L. Contributions to role-taking theory: A clinical study of self and role. *J. abnorm. soc. Psychol.*, 1952, 47, 117-125.
- SCHNECK, J. M. Spontaneous regression to an infant age level during self-hypnosis. *J. genet. Psychol.*, 1955, 86, 183-185.
- SEARS, R. R. Survey of objective studies of psychoanalytic concepts. *Soc. Sci. Res. Coun. Bull.*, 1943, No. 51.
- SPIEGEL, H., SHOR, J., & FISHMAN, S. An hypnotic ablation technique for the study of personality development. *Psychosom. Med.*, 1945, 7, 273-278.
- SUTCLIFFE, J. P. "Credulous" and "sceptical" view of hypnotic phenomena. *Int. J. clin. exp. Hypnosis*, 1960, 8, 73-102.
- TRUE, R. M. Experimental control in hypnotic age regression states. *Science*, 1949, 110, 583-584.
- TRUE, R. M., & STEPHENSON, C. W. Controlled experiments correlating electroencephalogram, pulse, and plantar reflexes with hypnotic age regression and induced emotional states. *Personality*, 1951, 1, 252-263.
- YATES, A. J. The use of vocabulary in the measurement of intellectual deterioration: A review. *J. ment. Sci.*, 1956, 102, 409-440.
- YOUNG, P. C. Hypnotic regression: Fact or artifact? *J. abnorm. soc. Psychol.*, 1940, 35, 273-278.

(Received August 2, 1960)

CHARLES DANTA COMPANY, INC., MILWAUKEE, WISCONSIN

